Assessment of Fibroblast Growth Factor Receptor cancer-associated mutations and characterisation of endocrine signalling complexes

Harshnira Hitesh Patani

University College London (UCL)
Institute of Structural and Molecular Biology

Thesis submitted for the degree of Doctor of Philosophy

March 2018

Declaration

I, Harshnira Hitesh Patani confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Publications

Patani, H., Bunney, T. D., Thiyagarajan, N., Norman, R. A., Ogg, D., Breed, J., Ashford, P., Potterton, A., Edwards, M., Williams, S. V., Thomson, G. S., Pang, C. S., Knowles, M. A., Breeze, A. L., Orengo, C., Phillips, C. & Katan, M. 2016. Landscape of activating cancer mutations in FGFR kinases and their differential responses to inhibitors in clinical use. Oncotarget, **7**, 24252-68.

Abstract

Receptor tyrosine kinase (RTK) signalling is frequently deregulated in cancers, developmental syndromes and metabolic diseases. The fibroblast growth factor (FGF) receptor (FGFR) family of RTKs have pleiotropic roles in development, metabolism and tissue homeostasis. A decade of deep sequencing studies has led to the discovery of somatic cancer-associated alterations in FGFRs including point mutations, gene fusions and amplifications. Missense substitutions appear most frequently in FGFR3 and FGFR2, some of which have been implicated as oncogenic drivers. This has led to rapid progress in anti-FGFR therapies.

This work compares and analyses twenty five reported FGFR kinase domain (KD) variants using FGFR3 as a model. Kinase activity assays were developed to screen missense substitutions *in vitro* using purified FGFR3 KD proteins. It was found that hyperactive mutations did not necessarily occur frequently and conversely, frequently occurring mutation 'hotspots' were not necessarily kinase hyperactive. Cellular models of FGFR mutations suggested that hyperactive mutants activated signalling but did not always transform cells. Investigations into the ligand binding characteristics of a subset of activating mutants revealed mutation specific signatures of binding properties. However, there appeared to be no link between KD activity and stability. These results reinforce the importance of screening clinical tumours prior to treatment to provide better data-driven therapies for patients.

Endocrine FGFR signalling was also explored in this work. Klotho family proteins bind to specific FGFs and FGFRs to signal through endocrine signalling complexes regulating diverse metabolic activities. This work developed and optimised methods to express and purify components of these binary and ternary signalling complexes. Preliminary biophysical experiments were then performed to begin to understand the protein-protein interactions involved in complex formation. This will inform the rational design of therapies already in development for targeting deregulated endocrine FGFR signalling in various metabolic disorders and cancer.

Acknowledgements

Firstly, I thank my supervisor Matilda Katan for the excellent opportunity to undertake this PhD in her lab. I am grateful for all the scientific guidance, learning, teaching and technical training opportunities she has provided. I have learned so much and have come a long way over four years. Secondly, I thank my industrial supervisors Richard Norman, Derek Ogg and Chris Phillips, all of who supervised me at various stages of the project and provided insight and guidance into industrial aspects of science. I thank my thesis committee, Lazaros Foukas and Ivan Gout, for their comments and helpful discussions throughout the project. I also thank Paul Driscoll for providing me with the early opportunities without which I would not be completing my PhD today.

Next, I thank my colleagues in the Katan lab for their constant support and friendship in the day-to-day lab. In particular, I am grateful to Tom Bunney, who started work on the FGFR projects before I joined the lab; thank you for answering all my silly questions over the years, sharing various materials and generally for sharing your extensive FGFR knowledge. I also thank Nethaji Thiyagarajan for advice and collaboration on the computational aspects of structural biology. I thank Hans Koss and Louis Perdios for all our scientific discussions, coffee drinking and late night laughs; your friendship made my time in the lab so much fun. Thank you also to Benedetta Lombardi, Domenico Sanfelice, Aleks Rust and Nicole Hartig for helpful discussions. Outside of the lab, I am very grateful to Mina Edwards and Sarah Williams for teaching me tissue culture and for advice and collaboration on many cell-related projects. Thank you also to Maggie Knowles for providing me with support and a lovely working environment during my time in Leeds. Thank you to all the staff at AstraZeneca for their support and technical guidance during my visits to their sites in Macclesfield and Cambridge. I thank Christine Orengo, Paul Ashford and Millie Pang for collaboration on the bioinformatics of the FGFR project, and Andrew Potterton for his excellent work with inhibitors.

Finally, I thank my parents for a lifetime of love, support and belief in me, and for ensuring I have always had every opportunity to enable me to be here today. Thank you to my dear sister Hemalvi for her friendship and constant support, and for all our scientific discussions, even at the dinner table. I thank my friends Lauren Matthews and Will Chaplin for always looking out for me, especially in the final months of writing up. Lastly, I thank Luke Marris for being my rock and for constant support, encouragement and motivation. Thank you all for being ‘on call’ for me throughout the PhD.

I also acknowledge my studentship funding provided by UCL and AstraZeneca.

Table of Contents

List of Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations

1/K	Time constant
2C	FGFR3-KD^{C482AC582S}
AB	Acid box
ABL	Abelson
ACC2	Acetyl-Co-A carboxylase 2
ACP	Adenylyl methylene diphosphonate
ADAM	A disintegrin and metalloproteinase
ADP	Adenosine diphosphate
AKT	Protein kinase B
ALK	Anaplastic lymphoma kinase
ALL	Acute lymphocytic leukaemia
A-loop	Activation-loop
AMPKα	Adenosine monophosphate-activated protein kinase α
ATP	Adenosine triphosphate
AU	Arbitrary units
BAD	BCL-2-associated death promoter
BAT	Brown adipose tissue
BAX	BCL2-associated X
BC	Breast cancer
BCL2	B-cell lymphoma 2
BCR-ABL	Breakpoint cluster region-Abelson
BF	Brightfield
BLAST	Basic Local Alignment Search Tool
Brij-35	Polyethylene glycol dodecyl ether
BSA	Bovine serum albumin
CASP 9	Caspase 9
CC	Colorectal cancer
CD	Circular dichroism
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulphonate
CHAPSO	3-[(3-Cholamidopropyl]-dimethylammonio)-2-hydroxy-1-propanesulphonate
CHO	Chinese hamster ovary
C-loop	Catalytic-loop
CML	Chronic myeloid leukaemia
COSMIC	Catalogue of Somatic Mutations in Cancer
C-spine	Catalytic spine
CTD	C-terminal domain
Cymal-5	5-cyclohexyl-1-pentyl-β-D-maltoside
CYP7A1	Cholesterol 7 α-hydroxylase 1
D1, D2, D3	Immunoglobulin-like domain1, 2, 3
DAG	Diacylglycerol
DDM	*n*-dodecyl-β-D-maltoside
DM	*n*-decyl-β-D-maltopyranoside
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
E. caratovora	*Erwinia caratovora*
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra-acetic acid
eGFP	Enhanced GFP
EGFR	Epidermal growth factor receptor

EGTA	Ethylene glycol tetra-acetic acid
EMT	Epithelial-mesenchymal transition
EPK	Eukaryotic protein kinase
ER+	Oestrogen receptor positive
ERK	Extracellular signal–regulated kinase
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
FGFR	FGF receptor
FL	Full length
FOX O	Forkhead box class O
FRET	Förster resonance energy transfer
FRS2(α)	FGF receptor substrate 2 (alpha)
FXR	Farnesoid X receptor
GAB1	GRB2-associated binding protein 1
Gdn	Guanidine
GFP	Green fluorescent protein
GIST	Gastrointestinal stromal tumour
GMQE	Global Model Quality Estimation
GRB2	Growth factor receptor bound protein 2
GSH	Reduced glutathione
GSK-3β	Glycogen synthase kinase 3β
GSSG	Oxidised glutathione
HBS	Heparin binding site
HCC	Hepatocellular carcinoma
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HER2	Human epidermal growth factor receptor
His	Histidine
HRP	Horseradish peroxidase
HS	Heparan sulphate
HSA	Human serum albumin
IB	Inclusion body
ICGC	International Cancer Genome Consortium
Ig	Immunoglobulin
IGF1	Insulin-like growth factor 1
IP_3	Inositol-1,4,5-trisphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IS, I	Insoluble, soluble
ITB	Inoue transformation buffer
ITC	Isothermal titration calorimetry
JM	Juxtamembrane
K	Rate constant
k_{cat}	Turnover number
K_D	Equilibrium dissociation constant
KD	Kinase domain
KI	Kinase insert
KIT	Stem cell growth factor receptor
KL1, KL2	Klotho domain 1,2
KLA	Klotho alpha
KLB	Klotho beta
K_M	Michaelis constant
KRB	Kinase reaction buffer

LB	Lysogeny broth
LCTL	Lactase-like
LIC	Ligation independent cloning
Lys	Lysate
mAb	Monoclonal antibody
MAPK	Mitogen-activated protein kinase
MARCKS	Myristoylated alanine-rich C-kinase substrate
MD	Macular degeneration
Med	Medium
MEK	Mitogen-activated protein kinase kinase
MET	Hepatocyte growth factor receptor
MTC	Medullary thyroid cancer
MWCO	Molecular weight cut-off
mβe	Mouse-βKlotho-eGFP
mβTe6H	Mouse-βKlotho-TEV-eGFP-6His
N-, C-	Amino-, carboxy-
NFAT	Nuclear factor of activated T cells
NFDM	Non-fat dried milk
NR	Non reducing
NRMSD	Normalised root mean square deviation
NSCLC	Non-small cell lung cancer
NTA	Nitrilotriacetic acid
NTD	N-terminal domain
OD_{600}	Optical density at 600 nm
OG	Octyl glucoside
OSCC	Oral squamous cell carcinoma
p38	p38 mitogen-activated protein kinase
p53	Tumour protein p53
pAb	Polyclonal antibody
PAGE	Polyacrylamide gel electrophoresis
PC	Pancreatic cancer
PCR	Polymerase chain reaction
PDGFR	Platelet derived growth factor receptor
PDK	Phosphoinositide dependent kinase
PEG	Polyethylene glycol
PEI	Polyethylenimine
pelB	Pectate lyase B
PI3K	Phosphoinositide 3-kinase
PIP_2	Phosphatidylinositol-4,5-bisphosphate
PIPES	Piperazine-N,N′-bis-(2-ethanesulphonic acid)
PKC	Protein kinase C
PLCγ	Phospholipase C gamma
P-loop	ATP-binding loop
$polyE_4Y_1$	Poly-glutamate-tyrosine
PPARα/γ	Peroxisome proliferator-activated receptor α/γ
PPS	3-(1-pyridinio)-1-propanesulphonate
PRAS40	Proline-rich AKT substrate of 40 kDa
pRB	Retinoblastoma protein
PTB	Protein tyrosine binding
PTH	Parathyroid hormone
PVDF	Polyvinylidene fluoride
QMEAN	Qualitative Model Energy Analysis

RAF	Rapidly accelerated fibrosarcoma
Ras	Rat sarcoma GTPase
RCC	Renal cell carcinoma
RCSB PDB	Research Collaboratory for Structural Bioinformatics Protein Data Bank
R_{max}	Maximum response
RMSD	Room mean square deviation
RNA	Ribonucleic acid
RSK2	Ribosomal protein S6 kinase 2
R-spine	Regulatory spine
RTK	Receptor tyrosine kinase
RU	Response units
SAS	Sequence Annotated by Structure
SCD1	Stearoyl-CoA desaturase 1
SCLC	Small cell lung cancer
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SFM	Serum-free medium
SH2	SRC homology 2
SH3	SRC homology 3
sKlotho	Soluble Klotho
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SOS	Son of sevenless
SP	Signal peptide
SPR	Surface plasmon resonance
STAT	Signal transducer and activator of transcription
StrepII	Streptavidin II
TB	Terrific broth
TC	Thyroid cancer
TCEP	Tris-(2-carboxyethyl)-phosphine hydrochloride
TCGA	The Cancer Genome Consortium
TE	Tris EDTA
TEV	Tobacco etch virus
TIM-barrel	Triosephosphateisomerase-barrel
TM	Transmembrane
Tris	Tris-(hydroxymethyl)-aminomethane
UPLC	Ultra performance liquid chromatography
UV	Ultraviolet
VEGFR	Vascular endothelial growth factor receptor
V_{max}	Maximum velocity
WAT	White adipose tissue
WCL	Whole cell lysate
WT	Wildtype

Chapter 1: Introduction

1.1 Introduction to receptor tyrosine kinases (RTKs)

The eukaryotic protein kinase (EPK) superfamily makes up one of the largest groups of homologous proteins. It constitutes 1.7% of the human genome with 500 members related by their function as enzymes that add a phosphoryl moiety to proteins in a reversible catalytic process known as the kinase reaction (Manning et al., 2002). In humans, the kinase activity of EPKs is mediated by their dynamic and flexible catalytic kinase domain (KD). This leads them to mediate most cellular signal transduction, regulating diverse cellular and metabolic functions through phosphorylation of specific amino acids in 30% of the human proteome. Despite diversity in sequence, structure, substrate specificities and regulation kinases share specific structural features of their catalytic core (Hanks and Hunter, 1995). Their amino acid-specific activities facilitate their grouping into two main subfamilies based on their target phosphorylation site: serine/threonine kinases, and tyrosine kinases. The tyrosine kinase family is further subdivided into receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases based on whether they possess, or lack, a transmembrane domain respectively. Fibroblast growth factor receptors (FGFRs) are members of the RTK subfamily that also includes other closely related members such as epidermal growth factor receptor (EGFR), platelet derived growth factor receptors (PDGFRs), vascular endothelial growth factor receptors (VEGFRs) and KIT.

RTKs are cell surface receptors that make up the second largest family of membrane-spanning proteins in metazoans. They mediate numerous physiological processes including cell growth, migration, survival, differentiation, development, metabolism and tissue homeostasis. The 58 known RTK members in humans are classified into 20 subfamilies based on sequence homology and domain organisation of family members (Lemmon and Schlessinger, 2010). All RTKs comprise an extracellular ligand binding domain linked to the intracellular kinase via a single-helical transmembrane domain (Figure 1.1). RTK activation and signalling occurs upon dimerisation and/or conformational changes induced by cognate ligand binding to the extracellular domain, which leads to *trans*-autophosphorylation of intracellular tyrosines and activation of the catalytic KD.

1.1.1 RTKs in cancer

More than half a century of RTK research has revealed that deregulated RTK signalling is responsible for several pathologies, from germline mutations causing developmental

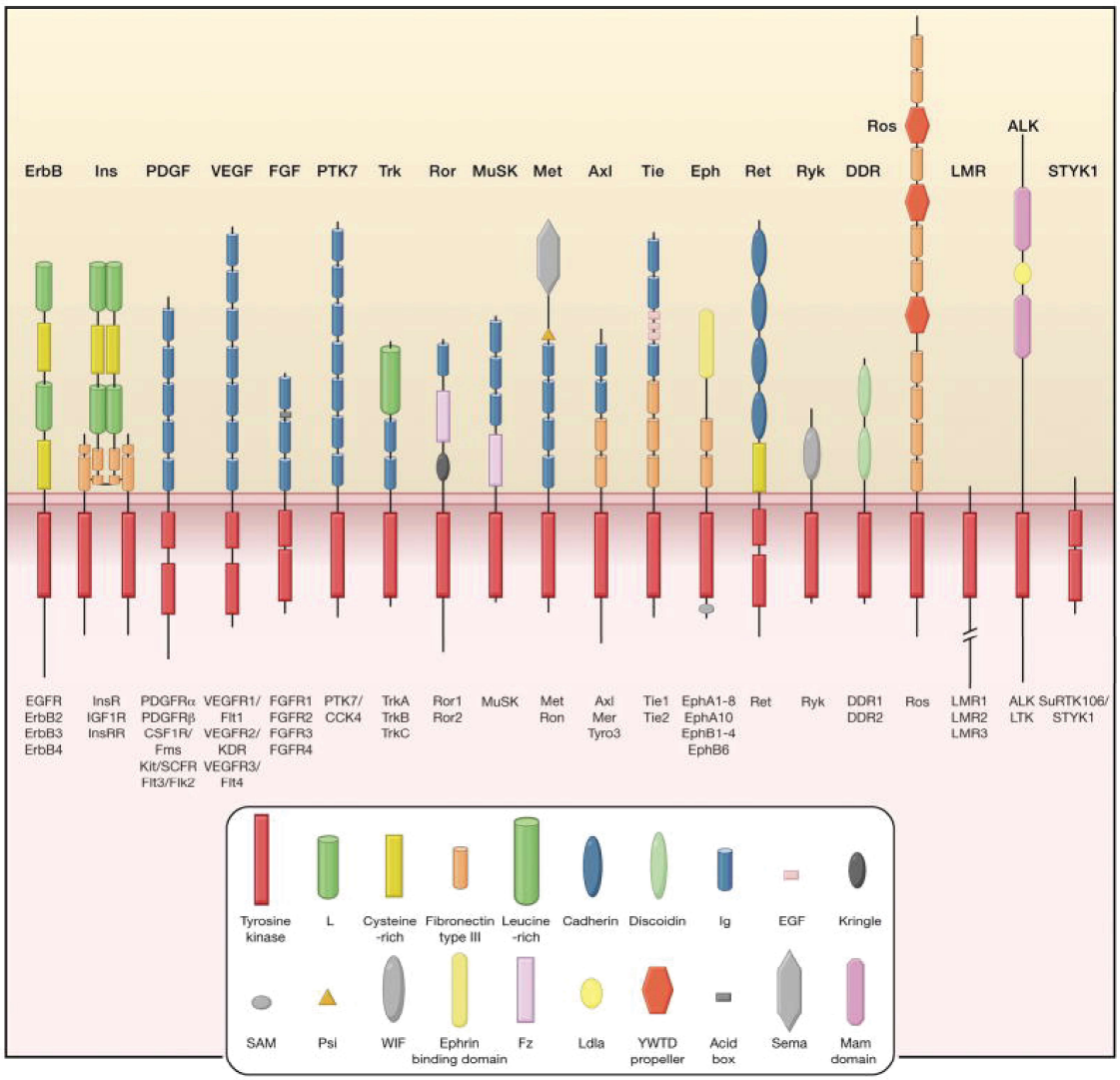


Figure 1.1 The human family of receptor tyrosine kinases (RTKs). The 58 members of the human family of RTKs are subdivided into 20 RTK subfamilies, each with one to five members per subfamily. The domain organisation of members within each subfamily is illustrated in schematic representation, showing the diversity of structural folds and domains of the extracellular domains within the RTK family. In comparison, the intracellular tyrosine kinase domain (the defining feature of the protein family, shown as red rectangles) is better conserved across subfamilies. Figure from (Lemmon and Schlessinger, 2010).

syndromes and somatic mutations contributing to oncogenesis (McDonell et al., 2015). At least ten RTK families have been observed to be mutated in congenital diseases, cancers, or both (Robertson et al., 2000). In addition there is substantial crosstalk between RTKs such that they are often co-activated in cancers (Xu and Huang, 2010) whereby two or more RTKs coordinate their signalling responses to evade regulatory cellular mechanisms that normally prevent tumour formation. Although RTKs were initially disregarded in drug development because of a lack of causative evidence for their carcinogenic roles alongside concerns over inhibitor specificity and on- and off-target toxicities, RTKs have been considered important targets in cancer for over a decade (Krause and Van Etten, 2005). Therefore, efforts to develop RTK targeted

therapies as treatments rapidly evolved leading to several RTK inhibitors progressing to clinical trials. The first inhibitor to be approved was Imatinib (2001) which specifically targets the non-receptor tyrosine kinase Breakpoint Cluster Region-Abelson (BCR-ABL) fusion protein found in chronic myeloid leukaemia (Druker et al., 1996). The success of this targeted therapy provided the necessary proof-of-concept evidence that fuelled the clinical progression of further anti-RTK clinical candidates and to date more than 20 have been approved to treat various malignancies (Table 1.1). Several of these therapies target relatively well-studied RTKs such as EGFR and VEGFR. Although many approved RTK inhibitors also exhibit variable levels of efficacy against other RTKs including FGFRs, due to their promiscuous nature and multi-kinase targets, therapies developed specifically against FGFRs are lagging behind other RTKs, and to date there are still no cancer therapies approved on the basis of FGFR targeting. FGFRs are known oncogenes, estimated to be mutated in about 7% of solid tumours of all cancer types (Guagnano et al., 2012, Helsten et al., 2016) and in haematological cancers, and remain important clinical targets. However, to date the efficacy of FGFR targeting in the clinic has been variable despite pre-clinical predictions of FGFRs as good therapeutic targets in all types of cancers (Babina and Turner, 2017). Additionally, the multi-kinase promiscuity of many approved RTK inhibitors causes on- and off-target toxicities of varying levels of seriousness in the clinic (Shah et al., 2013a, Shah et al., 2013b). In particular, effects of activating driver mutations on clinical efficacy of chemotherapies must not be underestimated. This is less well-studied and might account in part for the differences in patient sensitivity and response to targeted therapies observed in the clinic. Therefore, there is an urgent need for pre-clinical studies aimed at deepening our understanding of the biology, physiology and pathophysiology of oncogenic mutant RTKs. This will inform patient selection during clinical trials, accelerate the progression of targeted therapies (with fewer side-effects) up the clinical ladder and provide more efficacious therapeutic solutions for subsets of cancers, particularly those that warrant as yet insufficient clinical attention.

1.1.2 Acquired resistance to RTK inhibitors in the clinic

Another phenomenon which currently limits the efficacy of RTK inhibitors in the clinic is the emergence of resistance mutations in tumours in response to targeted therapies. Resistance mutations have been observed across kinases as a tumour relapse mechanism in response to ATP-competitive small molecule inhibitors, first observed in the clinic in the BCR-ABL kinase following Imatinib treatment (Gorre et al., 2001, Shah et al., 2002). Point mutations commonly occur at the ‘gatekeeper’ residue in the ATP binding pocket, hindering inhibitor binding through introduction of a sterically and chemically distinct residue (T315I in BCR-ABL) to the original gatekeeper residue.

Table 1.1 Current small molecule and monoclonal antibody inhibitors approved for the treatment of human cancers based on receptor tyrosine kinase (RTK) targeting.

Inhibitor (brand name)	Main RTK Targets	Cancer type	Year of first approval (US)
Afatinib (Gilotrif)	EGFR, HER2	NSCLC	2013
Alectinib (Alecensa)	ALK	NSCLC	2015
Axitinib (Inlyta)	VEGFR1-3	RCC	2012
Cabozantinib (Cometriq, Cabometyx)	MET, VEGFR2	MTC, RCC	2012
Ceritinib (Zykadia)	ALK	NSCLC	2014
Cetuximab (Erbitux)	EGFR	EGFR+ CC, NSCLC	2009
Crizotinib (Xalkori)	ALK	NSCLC	2011
Erlotinib (Tarceva)	EGFR, HER1	NSCLC, PC	2004
Gefitinib (Iressa)	EGFR	NSCLC	2009
Imatinib (Gleevec)	KIT, PDGFR	CML, GIST	2001
Lapatinib (Tykerb)	EGFR, HER2	BC, HER2+	2007
Lenvatinib (Lenvima)	VEGFR2-3	TC, RCC	2015/16
Osimertinib (Tagrisso)	EGFR	NSCLC, refractory	2015
Palbociclib (Ibrance)	ER+, HER2	BC, HER2-	2015
Panitumumab (Vectibix)	EGFR	EGFR+ CC	2006
Pazopanib (Votrient)	VEGFR1-3	RCC	2009
Pegaptanib (Macugen)*	VEGFR1-3	MD*	2004
Ponatinib (Iclusig)	PDGFR, VEGFR2	CML, ALL	2013
Regorafenib (Stivarga)	VEGFR1-3, PDGFR	CC, GIST	2012
Sorafenib (Nexavar)	VEGFR1-3	RCC, HCC	2005
Sunitinib (Sutent)	PDGFR, KIT	CML, GIST, RCC	2006
Trastuzumab (Herceptin)	HER2	HER2+ BC, GIST	2010
Vandetanib (Caprelsa)	VEGFR2	MTC	2011

RTK Abbreviations: EGFR [epidermal growth factor receptor]; HER [human epidermal growth factor receptor]; ALK [anaplastic lymphoma kinase]; VEGFR [vascular endothelial growth factor receptor]; MET [hepatocyte growth factor receptor]; KIT [stem cell growth factor receptor]; ER+ [oestrogen receptor positive]; PDGFR [platelet derived growth factor receptor].

Cancer Abbreviations: NSCLC [non small cell lung cancer]; RCC [renal cell carcinoma]; MTC [medullary thyroid cancer]; PC [pancreatic cancer]; CML [chronic myeloid leukaemia]; GIST [gastrointestinal stromal tumours]; BC [breast cancer]; TC [thyroid cancer]; ALL [acute lymphocytic leukaemia]; CC [colorectal cancer]; HCC [hepatocellular carcinoma].

* Inhibitor approved for treatment of the non-malignant disease, MD [macular degeneration]

Mutations at the corresponding gatekeeper threonine have been observed in relapsed tumours in response to inhibitors targeting various RTKs such as PDGFR (Cools et al., 2003), KIT (Tamborini et al., 2004), EGFR (Kobayashi et al., 2005, Pao et al., 2005) and anaplastic lymphoma kinase (ALK) (Choi et al., 2010). Therefore the effects of resistance mutations, including at non-gatekeeper residues, on clinical efficacy of RTK inhibitors must be anticipated as important limitations of targeted chemotherapy (Azam and Daley, 2006). More pre-clinical and clinical research is required to elucidate molecular mechanisms by which resistance mutations act to hinder inhibitor binding. This will facilitate the development of second generation RTK inhibitors to overcome the problem of clinical resistance by several distinct mechanisms.

FGFRs are strongly implicated oncogenes, commonly mutated in all types of cancers. Despite this, no approved clinical therapies exist, and additionally FGFRs have also been shown to acquire resistance mutations in a human tumour cell line in response to FGFR-targeted therapies (Chell et al., 2013). This pre-clinical study underscores the importance of a molecular understanding of therapeutic targeting in addition to *in vivo* and clinical studies in order to pre-empt possible 'escape' mechanisms of tumours in response to treatment. An understanding of molecular mechanisms by which other mutations might also affect inhibitor sensitivity, efficacy and ultimately FGFR pathophysiology in the clinic is vital to guide patient selection during clinical trials with the aim of a future personalised approach to cancer treatment. It is for these reasons that the biology, cancer mutations and signalling of FGFRs have been studied in this work from a molecular perspective, to contribute molecular detail to the emerging field of FGFR therapeutics which is still in its infancy. In the next section, the FGFR family of RTKs is introduced with respect to FGFR biology, physiology and signalling from a structural and biochemical perspective.

1.2 Introduction to FGFRs

The four main FGFR paralogues, FGFR1-4, comprise a modular structure typical of RTKs consisting of an extracellular region, a single-pass transmembrane domain and an intracellular region which contains the tyrosine KD. In normal physiology, signals are transduced through FGFRs in response to the extracellular domain binding to ligand fibroblast growth factors (FGFs) in the presence of specific proteoglycan co-factors or protein co-receptors. However, due to the involvement of FGFR signalling in numerous aspects of physiology, the consequences of pathway malfunction and deregulated FGFR signalling lead to the development of several diseases such as cancers, metabolic diseases and developmental disorders (Wilkie, 2005, Beenken and Mohammadi, 2009, Turner and Grose, 2010, Itoh and Ornitz, 2011, Carter et al., 2015).

In the next sections, the features of FGFs, FGFRs and their resultant ternary signalling complexes with co-factors or co-receptors that arbitrate successful signal transduction will be discussed. There will be a focus on the structural features of FGF and FGFR proteins that allow them to perform their specific functions, as well as on the molecular mechanisms that intricately activate and regulate FGF-FGFR signalling. The precise mechanisms behind deregulated FGFR signalling and the ensuing pathophysiological implications will be presented. Finally, the current breadth of therapeutic strategies in clinical development for targeting deregulated FGFR signalling will be summarised.

1.3 Overview of FGFR signalling

The FGF family of FGFR ligands comprises 22 members in humans, grouped into seven subfamilies based on their phylogeny, sequence homology and structure (Figure 1.2) (Itoh and Ornitz, 2004). Five of these subfamilies (comprising 15 FGFs) function as paracrine-acting 'canonical' ligands, while each of the other subfamilies (comprising three and four FGFs respectively) function as either intracellular-acting factors or endocrine-acting hormones respectively (Itoh and Ornitz, 2011). Canonical paracrine and hormonal endocrine FGFs are secreted proteins that bind to FGFRs at the cell surface. Paracrine FGFs form functional signalling complexes with FGFRs in the presence of heparan sulphate (HS) proteoglycan moieties at the plasma membrane (detailed in section 1.5.6), while endocrine FGFs bind with high affinity to FGFRs only in the presence of co-receptor proteins of the Klotho family and in a HS independent manner (discussed in detail in Chapter 4). Paracrine FGFs play vital roles in embryonic development and tissue homeostasis in adults, reviewed in (Ornitz and Itoh, 2015), while endocrine FGFs regulate diverse aspects of metabolism (reported in detail in Chapter 4). In contrast, although intracellular FGFs share structural homology to canonical FGFs, they are not secreted and do not bind to FGFRs (Olsen et al., 2003) and have FGFR-independent intracrine roles in the regulation of voltage gated sodium channels (Goldfarb et al., 2007). For these reasons they will not be discussed further.

The classical FGFR extracellular region comprises three immunoglobulin (Ig) fold-like domains (D1, D2 and D3) separated by short linker regions, linked via a single-pass type I transmembrane helix to the intracellular region which contains a juxtamembrane region and a 'split' tyrosine kinase catalytic domain (Figure 1.3). The four *FGFR* genes (*FGFR1-4*) are alternatively spliced to give rise to approximately 50 different FGFR isoforms, of which the most important splicing events affect the extracellular D3 domain of FGFR1-3 giving rise to canonical 'c' or alternative 'b' isoforms (detailed in section 1.4). FGF ligands signal through these splice isoforms in tissue specific pathways and the downstream

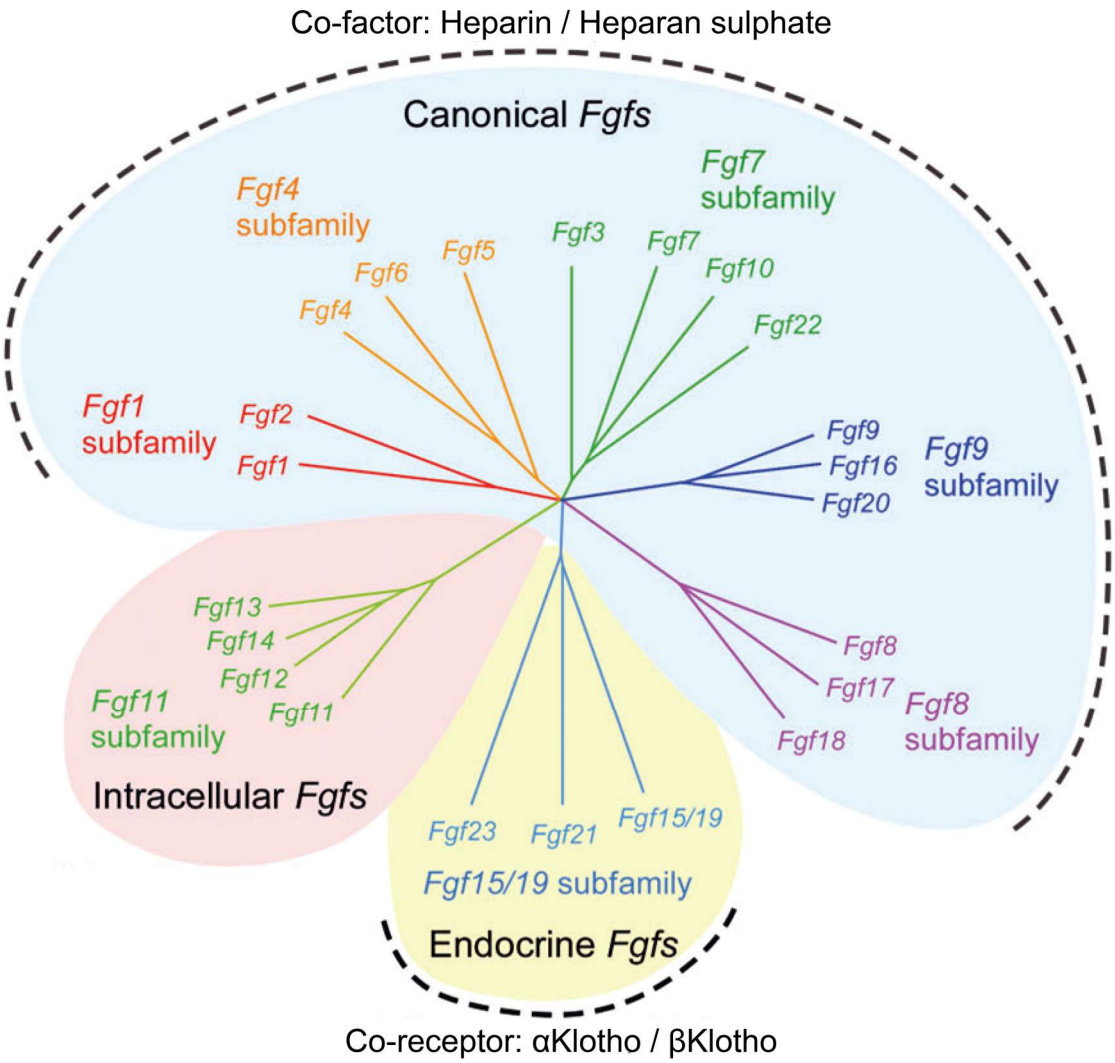


Figure 1.2 Phylogenetic classification of the human *FGF* gene family. The 22 *FGF* genes are grouped into seven subfamilies based on sequence and structure-based analyses, with two to four members per subfamily. The evolutionary distance between genes is proportionally represented by the lengths of the branches. The 15 'canonical' paracrine-acting FGFs make up five subfamilies and require heparin or heparan sulphate moieties as co-factors for signalling through FGFRs. The endocrine subfamily members act as hormones and require a Klotho family co-receptor for signalling. The 'intracrine' FGF subfamily do not bind or signal through FGFRs and act intracellularly as co-factors for other molecules. Figure from (Ornitz and Itoh, 2015).

molecular and physiological outcome of signalling depends on the particular FGF, FGFR and co-factor/co-receptor composition of the context-dependent ternary signalling complex. In general, paracrine and endocrine FGFs bind to the extracellular domains of FGFRs in the presence of HS co-factors or Klotho co-receptors respectively, triggering lateral dimerisation and/or activation of FGFRs at the cell surface and *trans*-autophosphorylation of the intracellular tyrosine KDs. The phosphorylated tyrosines on the activated KDs act as docking sites for the recruitment of substrate adaptor and effector proteins which also influence the specific response out of the myriad of possible downstream physiological outcomes (Figure 1.4). The two main intracellular FGFR

effectors are phospholipase C gamma (PLCγ) and the adaptor FGFR substrate 2 (FRS2) which bind to specific phosphorylated tyrosines on activated FGFRs and trigger distinct cellular signalling pathways (Figure 1.4). Spatial and temporal expression of FGFs and FGFRs (and Klotho co-receptors, discussed in Chapter 4) coupled with inherent sequence differences within FGFs and FGFR isoforms serve as key regulators of the outcome of FGFR signal transduction (Beenken and Mohammadi, 2009). The pleiotropic physiological roles of FGF signal transduction are presented in the next section, followed by the molecular determinants of FGFR signalling.

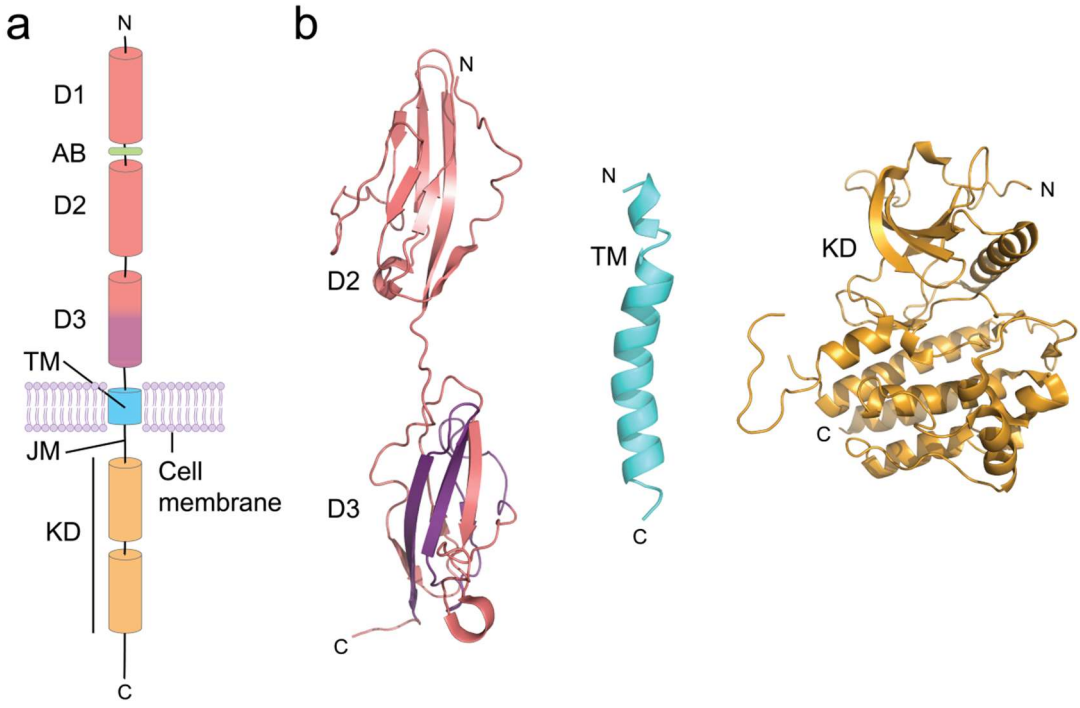


Figure 1.3 Domain organisation and representative structures of FGFRs. (a) Domain representation of the four members of the FGFR family of receptor tyrosine kinases (FGFR1-4). The extracellular region comprises three immunoglobulin-like domains (D1, D2 and D3) separated by short linker regions. The D1-D2 linker region contains a sequence of acidic residues termed the acid box (AB) (coloured green). The second half of the D3 domain (shaded purple) can be alternatively spliced to produce canonical 'c' or alternative 'b' splice isoforms. The ectodomain links via a single-pass transmembrane (TM) helix to the intracellular region which is comprised of a juxtamembrane (JM) segment and a 'split' catalytic tyrosine kinase domain (KD). **(b)** Structures of the domains in **a** are depicted in cartoon representation using representative crystal structures of FGFR3, coloured as in **a**. 'N' and 'C' label the amino- and carboxy-terminal ends of each structure. Shown are the D2-D3 extracellular region (left panel) [PDB: 1RY7] (Olsen et al., 2004), TM helix (middle panel) [PDB: 2LZL] (Bocharov et al., 2013) and both parts of the 'split' KD (right panel) [PDB: 4K33] (Huang et al., 2013b).

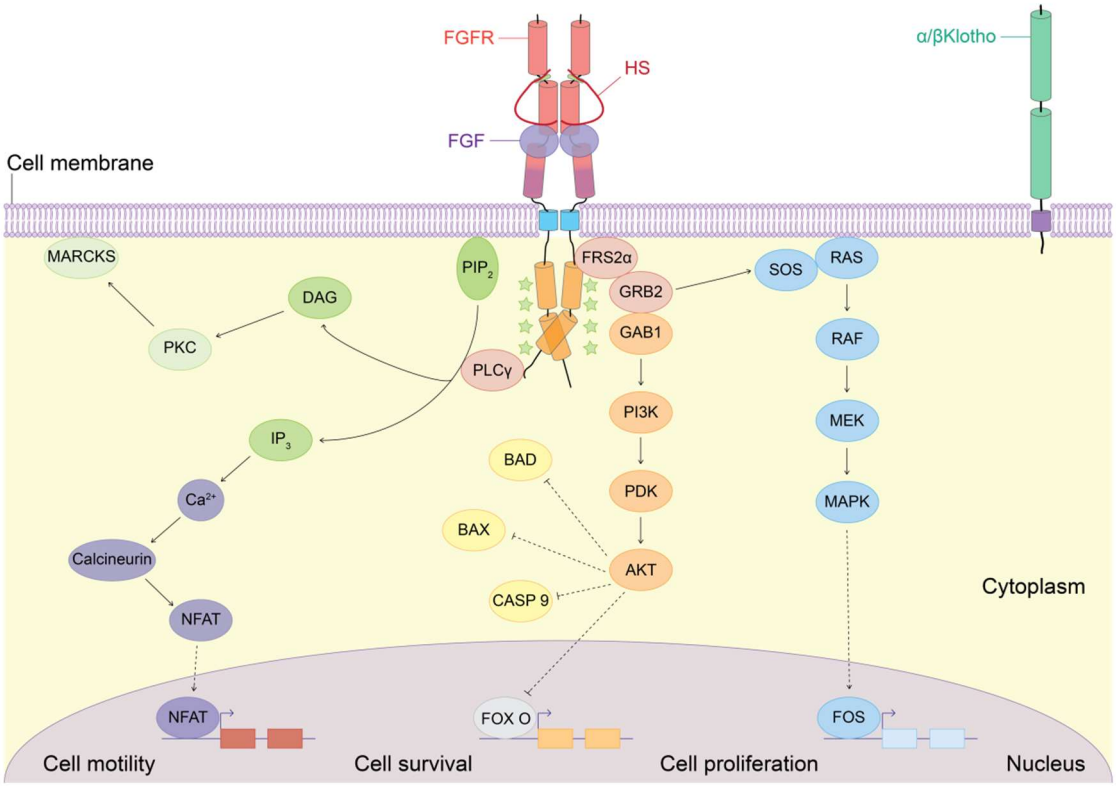


Figure 1.4 Overview of canonical paracrine FGFR signalling pathways. FGF ligands bind to FGFRs in the presence of heparan sulphate (HS) at the cell surface, inducing dimerisation of FGFRs and activation of the intracellular tyrosine kinase domains (KDs) through *trans*-autophosphorylation of KD tyrosines. Activated KDs phosphorylate tyrosines on downstream adaptor and effector proteins triggering signalling cascades that regulate a myriad of physiological outcomes. The two primary intracellular FGFR effectors are PLCγ and the FRS2α adaptor protein. Activated PLCγ hydrolyses PIP_2 to IP_3 and DAG. DAG activates PKC which phosphorylates its substrate MARCKS, while IP_3 stimulates the release of intracellular calcium, activating calcium-binding proteins such as calcineurin. Calcium-bound activated calcineurin induces NFAT to translocate to the nucleus, and MARCKS and nuclear NFAT activity then modulate cell motility. In comparison, FRS2α is constitutively bound to the juxtamembrane region of FGFRs where it also binds GRB2, serving as a signalling hub for the activation of distinct downstream pathways. In the first, GRB2 recruits the guanine nucleotide exchange factor SOS, which activates the Ras-RAF-MEK-MAPK cascade. Activated MAPK translocates to the nucleus and promotes cell proliferation through activation of immediate early gene transcription factors such as FOS. In the second, GRB2 recruits GAB1 which activates the PI3K-PDK-AKT pathway. Activated AKT inactivates the pro-apoptotic proteins BAD, BAX, Caspase 9 and the FOX O transcription factor, inducing a cell survival response. In a different cellular context for endocrine signalling, FGFRs associate with Klotho family proteins (instead of HS) and specific endocrine-acting FGFs to modulate aspects of metabolism through activation of the same downstream pathways (detailed in Chapter 4). Phospholipase C gamma (PLCγ), FGFR substrate 2α (FRS2α), phosphatidylinositol-4,5-bisphosphate (PIP_2), inositol-1,4,5-trisphosphate (IP_3), diacylglycerol (DAG), protein kinase C (PKC), myristoylated alanine-rich C-kinase substrate (MARCKS), nuclear factor of activated T cells (NFAT), growth factor receptor bound protein-2 (GRB2), son of sevenless (SOS), rat sarcoma GTPase (Ras), rapidly accelerated fibrosarcoma kinase (RAF), mitogen-activated protein kinases (MEK and MAPK), GRB2-associated binding protein 1 (GAB1), phosphoinositide-3-kinase (PI3K), phosphoinositide-dependent kinase (PDK), protein kinase B (PKB or AKT), BCL2-associated death promoter (BAD), BCL2-associated X protein (BAX), Caspase 9 (CASP 9), forkhead box class O (FOX O).

1.4 Physiological relevance of FGFR signalling

FGFs exert their wide-ranging and non-redundant biological effects in an organised, distinct FGFR-isoform, HS-specific and tissue-specific manner. Alternative splicing of FGFRs is largely tissue specific, with b and c splice isoforms expressed mostly in epithelial or mesenchymal tissues respectively (Miki et al., 1992, Orr-Urtreger et al., 1993, Goetz and Mohammadi, 2013). A wealth of research has elucidated the FGFR isoform specific complexes through which FGFs exert their effects *in vitro* and in cultured cells. It appears that FGFs secreted from mesenchymal tissues preferentially signal through epithelial FGFR b splice isoforms, while FGFs secreted from epithelial tissues interact predominantly with mesenchymal FGFR c splice isoforms (Finch et al., 1989, Ornitz et al., 1996, Zhang et al., 2006b). This results in reciprocal and directional paracrine signalling loops that coordinate embryonic development and adult tissue homeostasis in vertebrates while preventing illicit, pathophysiological autocrine signalling. Reciprocal loops of FGF signalling ubiquitously mediate all kinds of organogenesis, as reviewed in (Bottcher and Niehrs, 2005, Pownall and Isaacs, 2010). For example, mesenchymal expressed FGF10 signals through epithelial FGFR1b and FGFR2b splice isoforms while epithelial expressed FGF8 signals through mesenchymal FGFR2c forming reciprocal signalling loops necessary to initiate placenta and limb formation (Xu et al., 1998b). In a second example, epithelial FGF9 signalling through mesenchymal FGFR1c and FGFR2c causes mesenchymal expression of FGF10 which signals through epithelial FGFR2b to regulate caecal development (Zhang et al., 2006c). In addition, the formation of FGF7 and FGF10 gradients in the extracellular matrix (established by the variable affinities of these FGFs for HS, see section 1.5.1) modulate the extent and nature of branching morphogenesis in response to FGF threshold levels (Makarenkova et al., 2009). Conversely, endocrine FGF19 subfamily FGFs do not mediate this type of reciprocal signalling and organogenesis but control several aspects of metabolism in a HS independent and Klotho dependent manner (discussed in detail in Chapter 4).

A complete list of FGF-FGFR isoform specificities elucidated to date, summarising decades of work, are presented in Table 1.2. This summary demonstrates that members within FGF subfamilies mostly function through similar subsets of FGFR isoforms. FGFR paracrine signalling in normal physiology is therefore regulated by spatiotemporal and tissue specific expression of specific isoforms of FGFs and FGFRs (and HS – see section 1.5.6). In the next section, the structural and molecular basis for how the signalling outcome is determined from all possible physiological effects is described. This will be presented with a focus on the current understanding of specific structural features of

FGFs and FGFRs that govern FGF-FGFR-HS ternary complex formation, FGFR kinase activation and catalysis by the tyrosine KD.

Table 1.2 Receptor binding specificities of paracrine and endocrine FGFs for FGFR c and b splice isoforms.

FGF subfamily	FGF member	Co-factor / co-receptor	FGF receptor specificity
FGF1	FGF1	Heparan sulphate	All
	FGF2		FGFR 1c = 3c > 2c = 1b = 4
FGF4	FGF4		FGFR 1c = 2c > 3c = 4
	FGF5		
	FGF6		
FGF7	FGF3		FGFR 2b > 1b
	FGF7		
	FGF10		
	FGF22		
FGF8	FGF8		FGFR 3c = 4 > 2c = 1c
	FGF17		
	FGF18		
FGF9	FGF9		FGFR 3c = 2c > 1c = 3b >> 4
	FGF16		
	FGF20		
FGF19	FGF19	βKlotho	FGFR 1c = 4 > 2c > 3c
	FGF21		FGFR 1c > 3c > 2c >> 4
	FGF23	αKlotho	FGFR 1c = 4 > 3c > 2c

Compiled from: (Ornitz and Leder, 1992, Ornitz et al., 1996, Santos-Ocampo et al., 1996, Blunt et al., 1997, Ibrahimi et al., 2004, Itoh and Ornitz, 2004, Kurosu et al., 2006, Olsen et al., 2006, Urakawa et al., 2006, Zhang et al., 2006b, Ogawa et al., 2007, Kurosu et al., 2007, Suzuki et al., 2008, Harada et al., 2009, Ornitz and Itoh, 2015)

1.5 Structural determinants of FGF-FGFR-HS complex activation

Several models of FGFR activation have been proposed based on biochemical and structural studies. The two main models are the (1) diffusion based model and the (2) pre-formed dimer model (Figure 1.5). The diffusion based model has prevailed since the 1970s and describes a mechanism whereby RTK monomers move laterally within the cell membrane, occasionally stochastically colliding and resulting in basal activation in a diffusion dependent rate limiting process (Schlessinger et al., 1978). In this model,

dimerisation and FGFR activation is strictly ligand dependent. However, recent studies have demonstrated the presence of ligand- and density-independent phosphorylated FGFR dimers in cells (Comps-Agrar et al., 2015, Sarabipour and Hristova, 2016). In this pre-formed dimer model, inactive FGFR dimers (transmembrane domains and KDs in an 'inactive' configuration, non-permissive of autophosphorylation) are activated upon ligand binding by triggering a conformational switch to 'active' ligand-specific configurations of transmembrane helices, leading to correct juxtaposition of the KDs for activation. A similar mechanism has been proposed for other RTKs such as EGFR (Sarabipour, 2017). It is likely that a combination of both mechanisms exists *in vivo*, in context dependent situations. In this section, the components of FGFR paracrine signalling complexes will be introduced, detailing the specific inter-component binary contacts that mediate FGFR dimerisation and activation of the ternary complex. The endocrine FGFR signalling complexes (where HS is replaced by a Klotho co-receptor) will be discussed in detail in Chapter 4.

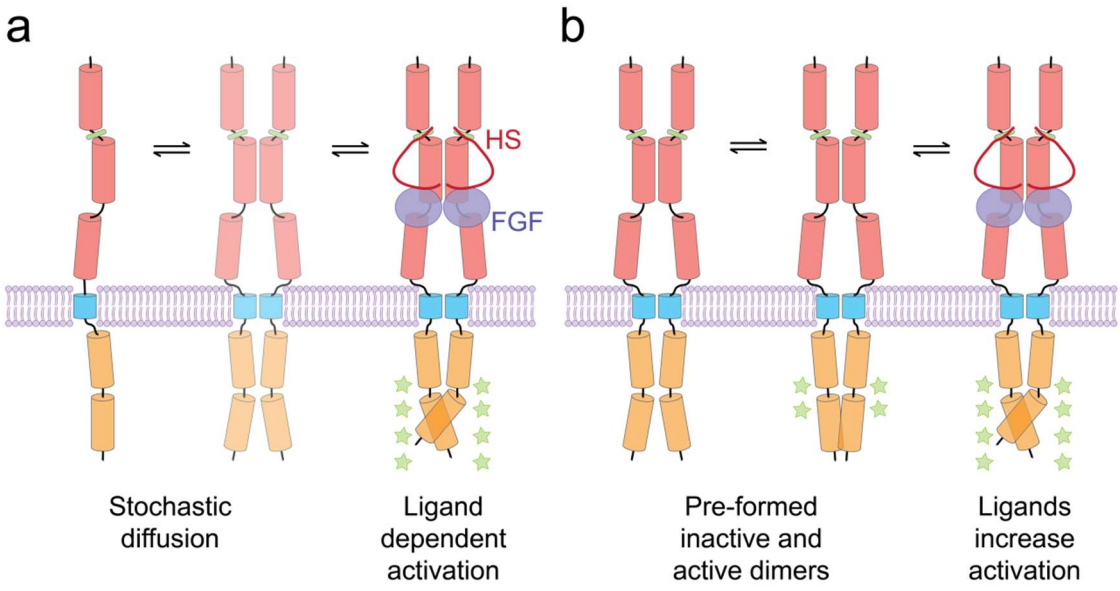


Figure 1.5 Models of FGFR activation. Two proposed working models for FGFR dimerisation and activation. **(a)** Diffusion based model. FGFR monomers can move laterally within the cell membrane in a diffusion-limited manner. Occasionally and stochastically they will collide, forming inactive dimers or resulting in basal activation, probably at a level too low for meaningful downstream signalling. The rate of activation is limited by diffusion (and thus concentration) of FGFR molecules within the membrane, and is strictly ligand dependent. **(b)** Pre-formed dimer model. In this model, FGFRs can exist in equilibrium between monomers and dimers at the cell surface, in a process not limited by diffusion or concentration of receptors at the cell surface. Some dimers are inactive, while some might exhibit low signalling activity depending on whether the juxtaposition of the kinase domains (KDs) is conducive for *trans*-autophosphorylation. In this scenario, ligand binding might serve to reorient the KDs to the correct orientation and/or conformation required for full activation, so signalling is partially ligand dependent.

1.5.1 Structural features of FGFs impart HS binding affinity and FGFR selectivity

The human family of FGFs vary in length between 150-300 amino acids (Itoh and Ornitz, 2004). The domain organisation of all FGFs and representative solved atomic structures of one FGF from each subfamily are shown in Figure 1.6. Most paracrine and the endocrine FGF subfamilies have an amino- (N-)terminal cleavable secretion signal peptide, while the FGF9 subfamily instead has a bipartite uncleaved secretion signal sequence (Figure 1.6a).

All FGFs comprise a 'core' homology domain which adopts a β-trefoil fold composed of approximately 120-125 amino acids. This comprises 12 anti-parallel β-strands (β1-β12) in paracrine FGFs. The β10-β12 region (encompassing the β11 strand), as well as other residues from regions such as the β1-β2 loop, form the primary site that gives FGFs the ability to bind HS (Figures 1.6a and 1.7a), the necessary proteoglycan cofactor for paracrine FGFR signalling as first shown for FGF1 and FGF2 (Thornton et al., 1983, Shing et al., 1984, Rapraeger et al., 1991, Ornitz and Leder, 1992, Ornitz et al., 1992). Indeed, even in this early work, it was shown that FGF1 and FGF2 had distinct binding affinities for HS (Lobb and Fett, 1984). Despite having a common structural topology, the HS binding site is sequence diverse across FGFs which results in the different binding affinities of FGFs for HS (Goetz et al., 2007, Makarenkova et al., 2009, Xu et al., 2012). All paracrine FGFs bind HS directly with high or moderate affinity and as a result, HS binding can alter FGF signalling capacity by altering conformation, enhancing thermal stability and protecting FGFs from proteolysis and inactivation (Gospodarowicz and Cheng, 1986, Saksela et al., 1988, Asada et al., 2009, Xu et al., 2012). In addition, the length and sulphation pattern of spatiotemporally expressed HS proteoglycans also determines the extent of FGF binding and signalling, detailed in section 1.5.6. HS binding ultimately controls the extent of diffusion of paracrine FGFs away from their site of secretion, allowing them to act locally (Goetz et al., 2007) and to form gradients that confer threshold-dependent signalling such as in developmental patterning and morphogenesis (Makarenkova et al., 2009). FGF mutations that affect HS binding can modulate their diffusion in the extracellular matrix, for example as shown for FGF9 (Harada et al., 2009, Kalinina et al., 2009).

In endocrine FGFs the β11 strand and its GxxxxGxx(T/S) motif (mostly conserved in paracrine FGFs) is missing and replaced by a helix at this position, resulting in loss of HS binding affinity for this FGF subfamily (Figure 1.7b). Additionally, endocrine FGF19 and FGF23 have two and one disulphide bonds respectively (Goetz et al., 2007, Harmer et al., 2004) and one of these disulphide linkages appears to be conserved across this

FGF subfamily. These key differences between paracrine and endocrine FGFs determine their divergent biological functions (detailed above and in Chapter 4).

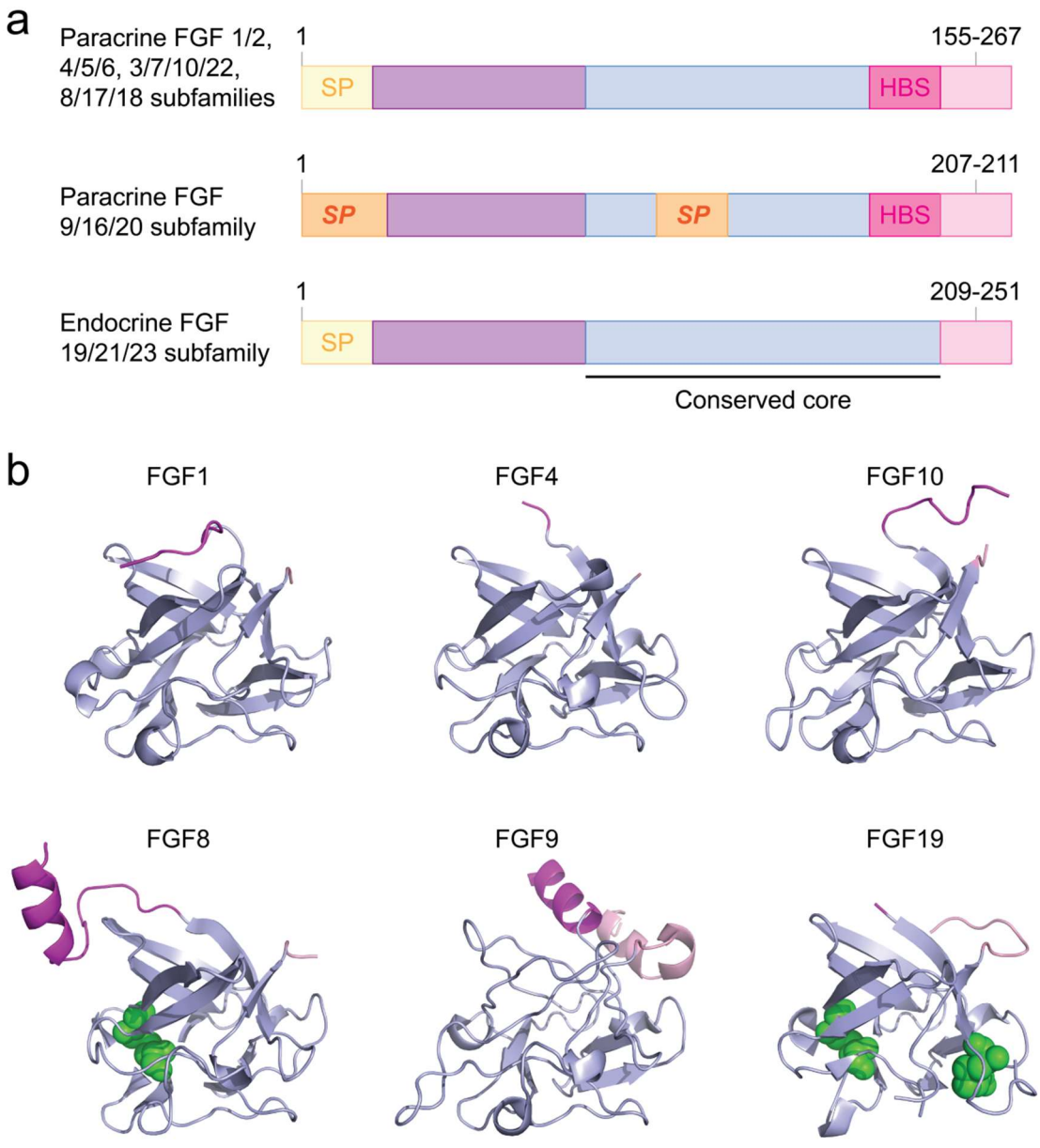


Figure 1.6 Domain organisation and representative structures of FGFs. (a) Domain representation of the paracrine and endocrine FGF subfamily proteins. All FGFs comprise an N-terminus (purple), a conserved 'core' (blue) and a C-terminal tail of variable length (light pink). They also comprise a cleavable N-terminal secretion signal peptide (SP, yellow) or a bipartite uncleaved secretion signal sequence (***SP***, orange). Paracrine FGFs also contain a heparin binding site (HBS, bright pink). **(b)** Structures of one member from each of the six FGF subfamilies depicted in cartoon representation and shown in the same orientation. N- and C-termini and the conserved 'core' are coloured purple, blue and light pink as in **a**. The cysteine residues that form disulphide bonds are highlighted as green spheres. FGF1 [PDB: 1RG8] (Bernett et al., 2004), FGF4 [PDB: 1IJT] (Bellosta et al., 2001), FGF10 [PDB: 1NUN] (Yeh et al., 2003), FGF8 [PDB:2FDB] (Olsen et al., 2006), FGF9 [PDB: 1IHK] (Plotnikov et al., 2001), FGF19 [PDB: 1PWA] (Goetz et al., 2007).

The FGF β-trefoil core is flanked by sequence diverse N- and carboxy- (C-)terminal regions (highlighted in Figure 1.6b) that confer sequence specific functional properties to FGF subfamilies. For example, the N-terminus of FGF8 is alternatively spliced to produce several splice isoforms with distinct FGFR binding specificities and biological functions (Crossley and Martin, 1995, Gemel et al., 1996, Blunt et al., 1997, Olsen et al., 2006). The N- and C-termini of FGF9 contribute to its unique ability to reversibly homodimerise (unlike other FGFs which only form monomers); part of the FGFR binding site is occluded in the homodimer interface and so the FGF9 dimer autoinhibits FGFR binding, modulating signalling by controlling the availability of FGFR-binding-competent FGF9 monomers (Harada et al., 2009, Kalinina et al., 2009, Plotnikov et al., 2001, Liu et al., 2017). Site-specific proteolytic cleavage of the C-terminal tail of FGF23 results in its inactivation, implicating the C-terminal tail in activation of FGF23 (Shimada et al., 2002, White et al., 2001). In fact, the C-terminal tails of all endocrine FGFs form the Klotho-FGFR binding sites, conferring the ability to form specific ternary complexes through which they can exert their biological functions (detailed in Chapter 4).

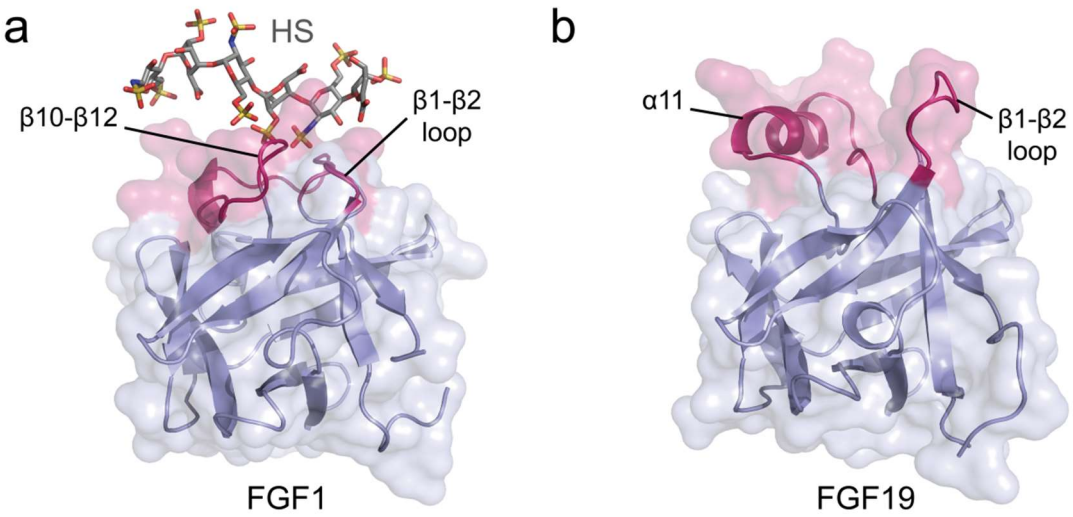


Figure 1.7 Comparison of the heparin binding site (HBS) on paracrine and endocrine FGFs. Structures of **(a)** paracrine FGF1 bound to a heparin hexasaccharide [PDB: 3OJV] (Beenken et al., 2012) and **(b)** endocrine FGF19 [PDB: 2P23] (Goetz et al., 2007) are shown in the same orientation in cartoon representation on a surface and coloured blue, with the heparin moiety shown as sticks and coloured grey and by atomic element with nitrogen atoms in navy blue, oxygen in red and sulphur in yellow. The HBS of each FGF is coloured bright pink, and the contributing β1-β2 loop and the β10-β12 regions are labelled. The pink region includes the GxxxxGxx(T/S) motif of the β11 strand in **a**. Note that the β1-β2 loop is extended and that β11 is replaced by α11 in **b** resulting in steric hindrance and partial loss of the HBS in FGF19, reducing its ability to bind heparin. Further details are provided in Chapter 4.

Structural studies with several FGFs have been useful in identifying specific sequences that confer FGFR binding specificity. Structural studies that included HS in the complex shed light on the fact that HS and FGFR binding sites on FGFs do not overlap (detailed in section 1.5.6). Several atomic structures of FGF-FGFR complexes have now been solved by crystallography, clearly demonstrating the diversity, yet similarity, between the FGF-FGFR interfaces. In general, the opposite end of the β-trefoil core to the HS binding site forms the FGFR binding region on paracrine FGFs (Mohammadi et al., 2005b). Visual examples of FGF-FGFR interfaces revealed through these crystal structures are shown in Figure 1.8. These highlight the diversity of the FGFR-binding interface of FGFs from different subfamilies (including the variable contributions of N- and C-terminal elements in determining binding specificity), as well as differences in FGFR-isoform specific conformations of key loops and other sequence elements on FGFRs towards different FGFs, which cannot be predicted by sequence analyses.

1.5.2 Structural basis by which FGFs bind to the FGFR ectodomain

Despite only sharing 55-72% sequence identity (Table 1.3a) the four FGFRs share overall structural features important for their function. The extracellular domain comprises Ig-like domains D1, D2 and D3 connected by flexible inter-domain linkers. A sequence of acidic S/D/E residues in the D1-D2 linker constitutes a feature unique to FGFRs called the 'acid box' (Figure 1.3a). The D1-D2 linker, including the acid box, and the terminal D1 domain, impart receptor *cis*-autoinhibition by competing with HS and FGFs for binding to the D2 domain, and exon skipping that excludes the D1 domain and/or the acid box in FGFR1-3 gives rise to FGFR isoforms with different affinities for FGFs (Wang et al., 1995, Shimizu et al., 2001, Olsen et al., 2004, Roghani and Moscatelli, 2007, Kalinina et al., 2012). Subtle sequence differences in the acid box of FGFR1-4 might translate to differences in the equilibrium between autoinhibited and non-autoinhibited conformational states, providing subtly different energy barriers for binding of different FGFs to their cognate receptor subtypes. Ultimately, the D1 domain and the D1-D2 linker are dispensable for paracrine ligand binding and the membrane proximal D2-D3 region on FGFRs is the necessary and sufficient binding site for paracrine FGFs, as shown by mutagenesis studies and numerous crystal structures (Johnson et al., 1990, Plotnikov et al., 1999, Stauber et al., 2000, Olsen et al., 2004, Mohammadi et al., 2005b). Therefore, FGFRs alternatively spliced to include or exclude the D1 domain are referred to as canonical 'α' (D1-D3) or alternative 'β' (D2-D3) splice isoforms respectively and have different affinities for paracrine FGFs. Additionally, as mentioned previously, alternative splicing of the FGFR1-3 D3 domain gives rise to b and c splice isoforms differing in the C-terminal region of the D3 domain that confers FGF binding specificity (Miki et al., 1992, Werner et al., 1992, Chellaiah et al., 1994, Yeh et al., 2003). The

structural elements that contribute to the FGF-FGFR interface differ slightly between b and c splice isoforms, but subtle differences between isoform specific residues are the key determinants of FGF binding specificity (Figure 1.8). The conformation and sequence of the βC'-βE loop in the D3 domain (the primary alternatively spliced region) is considered the key determinant of FGF binding specificity for FGFRs (Olsen et al., 2004), although FGF8 also interacts with a different region (the βF-βG loop, Figure 1.8c) (Olsen et al., 2006). Therefore the D3 domain, its alternative splicing and the relative D2 and D3 domain orientations together dictate FGF-FGFR binding specificity.

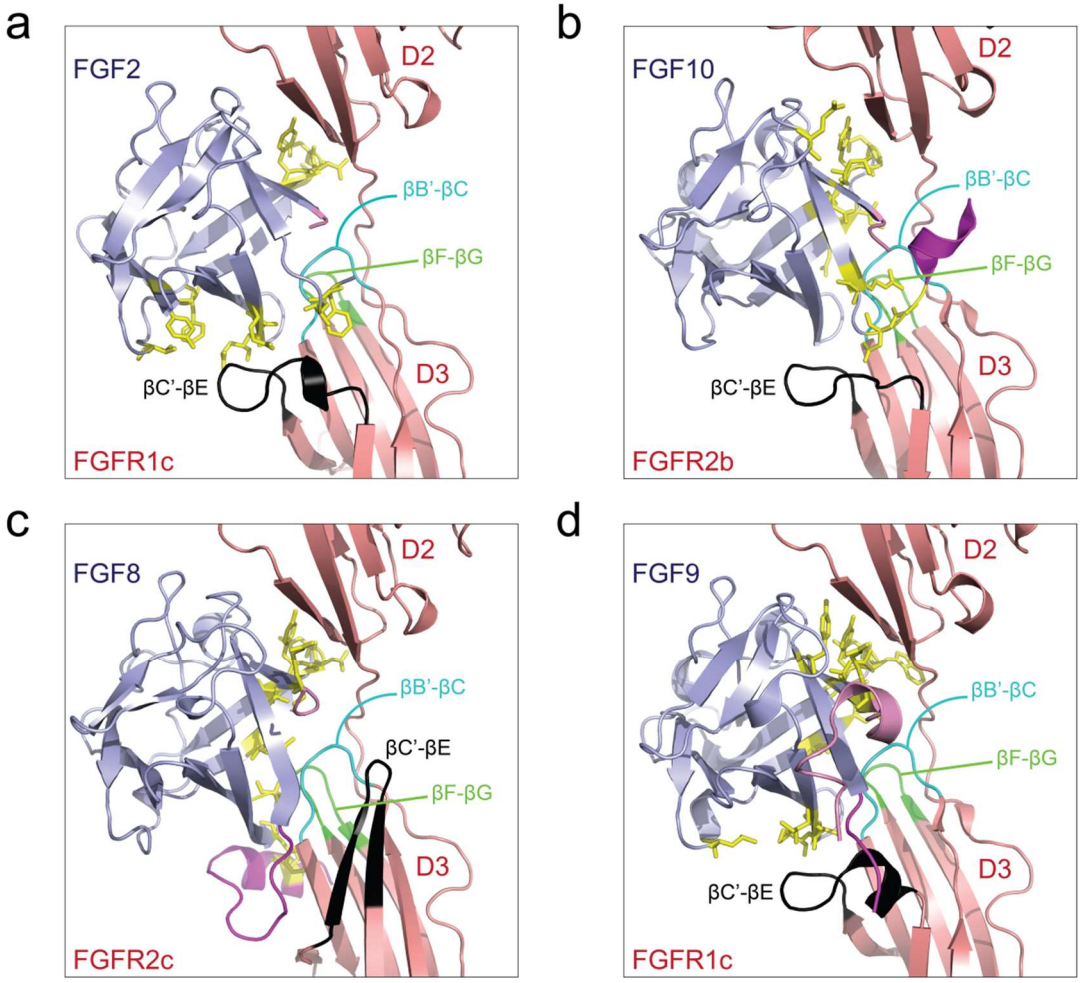


Figure 1.8 Structural basis for the variable receptor binding specificities of different FGFs and FGFR splice isoforms. Structures of different complexes of FGFs and FGFR D2-D3 domains of **(a)** FGF2-FGFR1c [PDB: 1FQ9] (Schlessinger et al., 2000), **(b)** FGF10-FGFR2b [PDB: 1NUN] (Yeh et al., 2003), **(c)** FGF8-FGFR2c [PDB: 2FDB] (Olsen et al., 2006) and **(d)** FGF9-FGFR1c [PDB: 5W59] (Liu et al., 2017) showing the binary 1:1 FGF-FGFR interfaces in cartoon representation and in the same orientation. FGFs and FGFR D2-D3 are coloured light blue and salmon respectively, as in Figures 1.3, 1.6 and 1.7. Additionally, the βB'-βC (cyan), βC'-βE (black) and βF-βG (green) loops that determine receptor binding specificity are highlighted and labelled on the FGFRs, and FGF residues at the FGFR interfaces are shown as sticks and coloured yellow (but are not labelled for clarity). Note the relatively invariant positions of the βB'-βC and βF-βG loops at the interface, while the position and/or ordering of the βC'-βE loop, and the relative orientation of the D3 domain, depends on differences in the N- and C-termini of the binding FGFs coupled with differences between FGFR isoforms.

Table 1.3 Percentage sequence **(a)** identity and **(b)** similarity between the full length proteins (light red) and kinase domains (light orange) of human FGFR1-4.

(a) Identity

FGFR4		79.9	75.8	76.8
FGFR3c	50.0		84.1	87.6
FGFR1c	55.7	63.5		87.6
FGFR2c	57.2	67.2	71.3	
	FGFR4	FGFR3c	FGFR1c	FGFR2c

(b) Similarity

FGFR4		84.5	83.1	82.8
FGFR3c	64.7		91.0	93.4
FGFR1c	61.7	69.9		91.0
FGFR2c	63.6	72.5	77.5	
	FGFR4	FGFR3c	FGFR1c	FGFR2c

1.5.3 The role of transmembrane and intracellular juxtamembrane regions

C-terminal to the extracellular region is the single helical transmembrane domain which links the FGFR ectodomain to the intracellular region comprising a juxtamembrane region and the tyrosine KD (Figure 1.3). Apart from putative dimer interfaces based on a single solution structure of the dimeric FGFR3 transmembrane helices (Bocharov et al., 2013), the mechanisms behind transmembrane helix dimerisation are unknown. It is hypothesised that residues of the transmembrane helix might not actually regulate dimerisation but interact to dimerise upon dimerisation of the ectodomains. A theoretical study also suggests a contribution of the C-terminal intracellular juxtamembrane region in dimerisation of the transmembrane helices (Peng et al., 2009). This region is a putative mediator of FGFR dimerisation because the precise formation of active, asymmetric dimers for the closely related EGFR KD is shown to be mediated by its intracellular juxtamembrane region through conformational coupling across the cell membrane (Zhang et al., 2006a, Jura et al., 2009, Red Brewer et al., 2009, Endres et al., 2013). Active asymmetric dimers of the FGFR KD have also been revealed by crystallography; in these dimers, one FGFR KD molecule serves as an enzyme, while the other acts as a substrate that is *trans*-phosphorylated by the enzyme-acting KD (Chen et al., 2008, Bae et al., 2010, Bae and Schlessinger, 2010). Although not experimentally shown for FGFRs, based on the EGFR models it is plausible that the intracellular juxtamembrane regions of FGFRs also juxtaposition the KDs upon ligand binding by transmembrane conformational coupling, forming active asymmetric dimers aligned for *trans*-autophosphorylation (Bocharov et al., 2013).

1.5.4 Architecture and activation of the FGFR KD

The structure of the catalytic tyrosine KD is highly conserved across EPKs. The FGFR KD is referred to as a 'split' KD, comprising a smaller N-lobe and a larger C-lobe. The two lobes behave as semi-rigid bodies in the inactive kinase and are connected by a short linker known as the 'hinge'. The N-lobe consists of five β-strands (β1-β5) forming a twisted β-sheet and a single α-helix (αC), while the C-lobe is primarily α-helical (αD-αE, αEF, αF-αI) with just two β-strands (β7-β8) (Figure 1.9a). The active site, where phosphoryl transfer occurs, lies in the cleft between the two lobes. Several other linear sequence elements are important for kinase function and activation (Figure 1.9a). These are the ATP binding loop (P-loop) and αC-helix in the N-lobe and the catalytic loop (C-loop), activation loop (A-loop), and kinase insert (KI) in the C-lobe. The triad of catalytic residues ('HRD' pocket) is located on the P-loop. The A-loop contains the conserved 'DFG' motif, which is involved in nucleotide and Mg^{2+} co-factor binding, and also contains two key tyrosine residues that undergo autophosphorylation during kinase activation (Figure 1.9a).

The KD acts as a molecular switch which is rapidly turned 'on' and 'off' in response to biological cues. The switch from active to inactive kinase is a transient, complex and highly regulated process resulting from dynamic conformational changes in the listed structural elements upon phosphorylation of the A-loop tyrosines (Figure 1.10). In the inactive KD the configuration of the A-loop, anchored by the αC-helix, in the 'closed' conformation autoinhibits activity in *cis* by hindering protein substrate binding, without affecting ATP binding to the active site (Mohammadi et al., 1996b). In addition, a triad of residues (non-consecutive in sequence) near the kinase hinge form an autoinhibitory network of hydrogen bonds termed the 'molecular brake' which maintains the KD in its inactive state (Chen et al., 2007b). Extracellular ligand stimulation of FGFRs leads to structural rearrangements of the ectodomains (see section 1.5.2) that translate to juxtaposition of the KDs in the correct asymmetric conformation required for *trans*-autophosphorylation of the essential A-loop tyrosines (Mohammadi et al., 1996a, Bae et al., 2010). This activates the KDs through 'release' of the molecular brake, rotation of the N-lobe towards the C-lobe, stabilisation of the 'open' configurations of the αC-helix and the A-loop, and reorientation of key catalytic residues (Figure 1.10). This promotes autophosphorylation of other intracellular tyrosine residues and permits substrate binding (Hubbard, 1999, Belov and Mohammadi, 2013, Chen et al., 2017).

In addition, there are two non-linear hydrophobic 'motifs' (composed of residues distant in sequence space) that are vital for kinase activation. These are the catalytic (C) and regulatory (R) spines, anchored to the N- and C-terminal ends of the αF-helix

respectively (Figure 1.9b). Dynamic assembly of the R-spine and its surrounding 'shell' residues into an ordered structure is dependent on the active A-loop configuration and is mandatory for kinase activation, while disassembly leads to kinase inactivation (Kornev et al., 2006, Meharena et al., 2013). Assembly of the C-spine during kinase activation is completed upon binding of the adenine ring of ATP (Kornev et al., 2008). These spines comprise residues vital for catalysis such as the H and F residues from the HRD pocket and DFG motif respectively (Figure 1.9b). Assembly of the spines stabilises the active KD conformation by connecting residues from the N- and C-lobes and securing the positions of catalytic residues, ATP and substrates for catalysis (Taylor and Kornev, 2011).

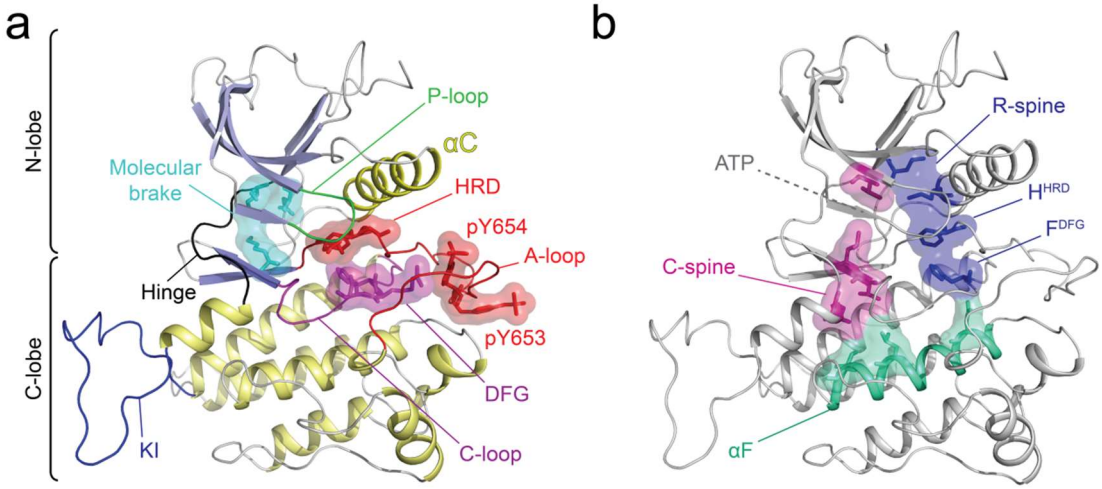


Figure 1.9 Architecture of the FGFR tyrosine kinase domain (KD). The linear and non-linear sequence features that are important for KD function and activation are shown on the structure of phosphorylated and active FGFR1 [PDB: 3GQI] (Bae et al., 2009) shown in cartoon representation. **(a)** The FGFR KD is 'split' into an N-lobe and a C-lobe connected by a flexible linker known as the kinase 'hinge' (black) which allows relative movement of the N- and C-lobes. The N-lobe comprises five β-strands that form a twisted β-sheet (β1-5, light blue) and a single helix (αC, yellow). The C-lobe has two β-strands (β7-8, blue) near the hinge region but is primarily α-helical (pale yellow) and comprises a loop unique to FGFRs known and the kinase insert (KI, dark blue). The active site lies in the cavity between the N- and C-lobes. The nucleotide binding loop (P-loop, green) from the N-lobe and the catalytic loop (C-loop, purple) and activation loop (A-loop, red) from the C-lobe contribute to this active site, and specific linear sequence motifs in the C-loop and A-loop (HRD and DFG respectively, shown as sticks with a surface on each loop) provide the vital residues for ATP and Mg^{2+} co-factor binding and phosphoryl transfer. Also shown in sticks against a surface (and labelled) are the crucial 'twin' tyrosines on the A-loop that undergo *trans*-autophosphorylation for full activation of the KD. Three residues (non-linear in sequence) in the hinge region which form an autoinhibitory network of interactions in the *inactive* kinase ('molecular brake', cyan) are highlighted as sticks against a surface. **(b)** Residues from both lobes of the KD forming two hydrophobic 'motifs' (non-linear in sequence) whose conformations are crucial for kinase activation, form the regulatory and catalytic R-spine and C-spine (shown as sticks against a surface and coloured dark blue and bright pink respectively). Assembly of the spines is crucial for kinase activation, as shown on this structure of an active FGFR KD. Full assembly of the C-spine is completed with the adenine ring of ATP, which slots into the labelled cavity. The spines are 'anchored' to the αF-helix in the C-lobe (aquamarine) with the 'base' residues of each spine contributed by this helix. The R-spine contains two key catalytic residues from the HRD and DFG motifs shown in **a**, and these are labelled.

The mechanisms involved in full activation of the KD are partially understood, and key questions remain as to the extent, duration and order of autophosphorylation during activation. Autophosphorylation of all tyrosines on the FGFR1 KD has been examined and appears to follow distinct stages of kinetically driven sequential *trans*-autophosphorylation of the key 'twin' tyrosine residues in the A-loop and other tyrosines in the juxtamembrane, KI and C-terminal tail regions (Furdui et al., 2006, Lew et al., 2009). The sequential order of autophosphorylation of tyrosines in FGFR2 has also been studied and differs from the order of phosphorylation observed for FGFR1 (Chen et al., 2008). Apart from the A-loop tyrosines, other tyrosines in the FGFR KDs are not all conserved, suggesting paralogue specific differences in autophosphorylation between FGFR1-4 which remain to be studied. This might translate to differences in the extent of kinase activity upregulation and in the effector proteins recruited and subsequent activation of context dependent downstream signalling cascades.

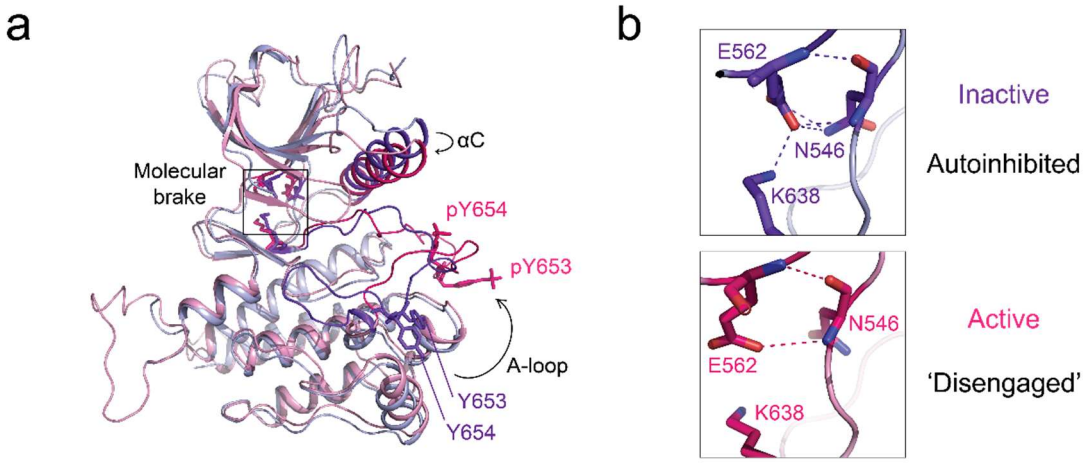


Figure 1.10 Switch from inactive to active conformation of the FGFR kinase domain (KD). **(a)** Overlay of FGFR1 KD inactive (light purple) [PDB: 1FGK] (Mohammadi et al., 1996b) and active (light pink) [PDB: 3GQI] (Bae et al., 2009) structures shown in cartoon representation, showing the global conformational differences between the two KD activation states. The features whose conformations are particularly important during the switch are highlighted in dark purple (inactive) and bright pink (active) respectively and the positions of key residues are shown as sticks. Inward rotation of the αC-helix and movement of the A-loop (and its 'twin' tyrosines, phosphorylated only in the active kinase) to the 'open' and active conformation is clearly visible. The autoinhibitory molecular brake residues in the kinase hinge region are boxed. **(b)** Close-up view of the boxed molecular brake residues in **a**, showing the network of hydrogen bonds (dashed lines) in the inactive kinase (upper panel) which is disrupted in the active kinase (lower panel).

1.5.5 Role of the C-terminal tail in regulating kinase activity

The C-terminal tail region (following the KD) has not been well studied with respect to its effect on kinase activity so little is known about any regulatory role this region plays in modulating dimerisation and/or kinase activation *in vivo*. This region contains several tyrosines that are phosphorylated during kinase activation and serve as docking sites for recruiting downstream effectors (Figure 1.4). None of the structures of the FGFR KDs solved to date include the C-terminal tail residues. However, one study has shown that the FGFR effector protein growth factor receptor bound protein-2 (GRB2) binds to a proline rich region on the C-terminal tail of FGFR2 in cells via its C-terminal SRC homology 3 (SH3) domain (Ahmed et al., 2010). A model was subsequently proposed whereby GRB2 dimers regulate FGFR2 activity by binding to this region on FGFR2 in the absence of ligand, forcing dimerisation of FGFR and 'basal' level phosphorylation of the A-loop, 'priming' the FGFR dimer for rapid activation upon ligand binding (Lin et al., 2012). However, GRB2 binding sterically inhibited phosphorylation of other tyrosines on FGFR and obstructed the recruitment of other downstream effectors, thus regulating FGFR signalling. Furthermore, upon stimulation with FGF, GRB2 was phosphorylated and released leaving the FGFR free to initiate full activation and signal transduction. This model supports the pre-formed dimer model of FGFR activation (Figure 1.5b) (Belov and Mohammadi, 2012).

1.5.6 Assembly of the signalling competent FGF-FGFR-HS ternary complex

The lack of structures of any full length FGFRs means a complete understanding of how all the domains are coupled across the membrane to transduce the intracellular signal based on extracellular ligand binding is not yet available. However, several crystal structures of FGF-FGFR or FGF-FGFR-HS solved using FGFR ectodomains, paired with several structures of the KD and biochemical experiments in cells using full length FGFRs, have been informative about some of the mechanisms behind complex formation. Dimerisation of FGFRs to form signalling competent FGF-FGFR-HS ternary complexes requires a combination of direct binary contacts between ligands and receptor molecules, namely (1) FGF-FGFR, (2) FGFR-FGFR, (3) FGF-HS and (4) FGFR-HS interactions.

Two early crystal structures of a dimer of FGF1:FGFR2 or FGF2:FGFR1 heterodimers proposed that FGF-FGFR interactions in the ternary complex had a molecular stoichiometry of 2:2 (Plotnikov et al., 1999, Stauber et al., 2000). They showed that primary contacts between the D2 domains of the two FGFRs in the complex, contacts within 1:1 FGF-FGFR heterodimers (at the D2-D3 domain), and secondary contacts between the FGF from one 1:1 heterodimer and FGFR from the other, contributed to

assembly of the complex. These studies also mapped a putative HS binding site to a region between the two FGFs that contacts both FGFs and the FGFR D2 domains. Concurrently, two structures of the FGF2-FGFR1-HS (Schlessinger et al., 2000) or FGF1-FGFR2-HS (Pellegrini et al., 2000) ternary complexes were also solved. Both structures confirmed the 2:2 stoichiometry of FGF-FGFR in each complex, but differed in the stoichiometry of HS. The Pellegrini model proposed an asymmetric 2:2:1 model in which a single HS molecule bound to a different positively charged FGF-FGFR interface, but this model is thought not to be biologically relevant. The Schlessinger model proposed a symmetric 2:2:2 complex of FGF-FGFR-HS with the HS forming several hydrogen bonds with the positively charged 'canyon' region between FGFs and the FGFR D2 domains. HS stabilises the complex by simultaneously interacting with both FGF and FGFR in the 1:1 heterodimer prior to homodimerisation of this 1:1:1 trimer to the 2:2:2 complex (Figure 1.11). The structure highlights the different contributions of FGF-FGFR, FGFR-FGFR, FGF-HS and FGFR-HS interactions cooperating to stabilise the FGFR dimer. This model is supported by the majority of experimental evidence to date, as reviewed in (Mohammadi et al., 2005a).

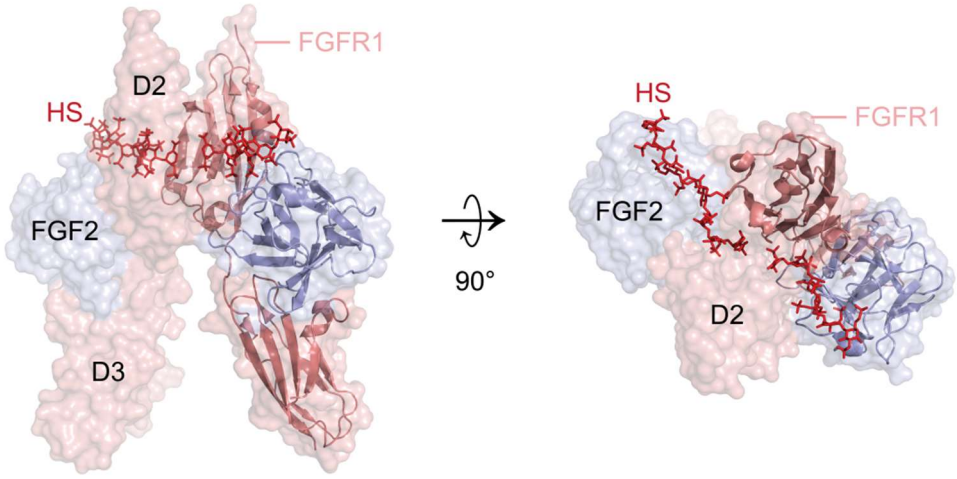


Figure 1.11 Structure of the extracellular portion of the ternary signalling complex. The crystal structure of the extracellular portion of a signalling competent 2:2:2 ternary complex [PDB: 1FQ9] (Schlessinger et al., 2000) comprising two molecules of each of FGF2, FGFR1 (D2-D3) and a heparin sulphate (HS) octasaccharide is shown. The two orientations are rotated by 90° relative to each other, showing the complex from the 'front' (left panel) and from 'above' (right panel). For clarity, only one FGF (light blue) and FGFR (salmon) molecule is shown in cartoon representation against the surface. Both HS molecules are shown as sticks and coloured red. The HS binds in the positively charged 'canyon' between FGF2 and the FGFR1 D2 domain where it contacts an FGF and FGFR molecule within each 1:1:1 FGF-FGFR-HS trimer in the 2:2:2 complex. Contacts between FGF2 and both FGFR molecules in the complex are clearly visible from 'above' (right panel) while contacts between the D2 domains of the FGFRs can be observed in both depicted orientations.

Although HS is required for high affinity binding of paracrine FGFs to FGFRs (Yayon et al., 1991), the minimal length of HS required for FGFR activation in cells was found to be an octasaccharide (Ornitz and Leder, 1992). However, *in vitro* experiments and the crystal structures above suggest that shorter HS chains might be sufficient for binding to the FGF-FGFR binary complex (Schlessinger et al., 2000). The observation that spatiotemporal differences in HS lengths and sulphation patterns direct signalling through different combinations of FGFs and FGFRs suggests that complex assembly through FGFs and FGFRs with different affinities for HS will be further fine-tuned by their ability to bind to tissue specific HS chains to form signalling complexes (Ostrovsky et al., 2002, Qu et al., 2011, Qu et al., 2012). Different HS binding affinities establish variable FGF diffusion radii within the extracellular matrix, manifesting as (reciprocal) FGF gradients that modulate their biological function and the extent of their response (Makarenkova et al., 2009). A final level of regulation of extracellular dimer formation depends on the extent of autoinhibition by the D1 domain and the acid box region, discussed in section 1.5.2. A shift in the equilibrium to more FGFRs in the non-inhibited 'open' D1 conformation will allow greater levels of ligand binding and activation, and vice versa.

The current premise is that simple dimerisation of FGFRs is insufficient for activation, and that the correct orientation of domains within the dimer is required to fully activate the kinase. Subtle FGF specific rotations of D2 and D3 domains of FGFR b and c splice isoforms upon FGF-FGFR binding might propagate to differences in the juxtaposition of intracellular KDs, modulating the signal (Yeh et al., 2003, Olsen et al., 2006). The intracellular FGFR-FGFR contacts are vital for full activation of the complex. As discussed in section 1.5.3, it is plausible that this is modulated through conformational coupling across the plasma membrane. In the case of the diffusion based model, this will be a direct consequence of extracellular ligand binding. In the case of the pre-formed dimer model, there might be a switch from an autoinhibitory, symmetric conformation of inactive KD dimers (Mohammadi et al., 1996b) to correctly juxtaposed asymmetric KD dimers that can *trans*-phosphorylate (Chen et al., 2008, Bae et al., 2010). The asymmetric FGFR1 and FGFR2 homodimers in these structures have very different interfaces between enzyme- and substrate-acting KDs, suggesting that sequence specific differences between the D3 domains (b and c splice isoforms) and KDs of FGFR1-4 will determine the nature of intracellular FGFR-FGFR contacts and tyrosine *trans*-phosphorylation, thus affecting activation and downstream signalling.

1.6 Activation of downstream FGFR signalling cascades

Upon FGFR activation, the two main downstream effectors that are activated through phosphorylation by the activated FGFRs are PLCγ and the adaptor protein FRS2 (Figure 1.4). PLCγ binds to a phospho-tyrosine residue towards the C-terminus of FGFRs via one of its SRC homology 2 (SH2) domains (Mohammadi et al., 1991, Bae et al., 2009). Tyrosine phosphorylation of PLCγ by FGFR activates PLCγ causing it to hydrolyse phosphatidylinositol-4,5-bisphosphate (PIP_2) to inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) at the plasma membrane (Carpenter and Ji, 1999). This activates protein kinase C (PKC) and the release of calcium from intracellular stores, affecting cell motility. In contrast, FRS2 binds to a phospho-tyrosine on the FGFR juxtamembrane region via its phospho-tyrosine binding (PTB) domain and, once activated by phosphorylation, serves as a hub for the recruitment of more downstream effectors through binding to GRB2 (Kouhara et al., 1997). GRB2 then activates the Ras-RAF-MEK-MAPK pathway through recruitment of the son of sevenless (SOS) guanine nucleotide exchange factor, and activates the PI3K-PDK-AKT pathway through recruitment of the GRB2-associated binding protein 1 (GAB1), leading to pro-survival signalling and proliferation (Figure 1.4). Other pathways can also be activated by FGFRs in different cellular contexts linked to cellular transformation, such as signal transducer and activator of transcription (STAT) signalling (Hart et al., 2000, Krejci et al., 2008, Wu et al., 2013) and ribosomal S6 kinase 2 (RSK2) signalling (Kang et al., 2009).

1.7 Higher order regulation of FGFR signalling

1.7.1 Cell surface FGFR internalisation by endocytosis

Ligand-induced internalisation of RTKs is believed to play an important role in the modulation of signalling at the cell surface (Sorkin and von Zastrow, 2009). Several reports have shown that ligand bound FGFRs can also be internalised via clathrin- and dynamin-dependent or independent routes, based on which receptor isoform is being internalised, and FGF-FGFR complexes are then either recycled to the plasma membrane or degraded in lysosomes (Haugsten et al., 2005, Sandilands et al., 2007, Haugsten et al., 2011). Ubiquitination of the receptor has been shown to be necessary for intracellular sorting of internalised FGFRs to lysosomes for degradation but not for endocytosis (Haugsten et al., 2008). Additionally, it has been suggested that receptor dimerisation might be sufficient, and kinase activity and ligand binding dispensable, for internalisation of FGFRs (Opalinski et al., 2017). This leads to the hypothesis that receptor internalisation might regulate ligand-independent signalling by removing FGFR dimers from the cell surface in the absence of ligand stimulation, dampening basal level

signalling. In addition, the duration, and therefore potency, of the transduced signal upon ligand binding might depend on the level of FGFR ubiquitination and the rate of FGF-FGFR complex recycling versus degradation. These open questions will require experimental testing to understand the involvement of endocytosis in the regulation of FGFR signalling.

1.7.2 Regulation of signalling via extracellular and intracellular docking proteins

The Klotho family of transmembrane FGFR co-receptor proteins convert the FGFR signalling mode from paracrine to endocrine in response to the endocrine FGF19 subfamily ligands (detailed in Chapter 4). An FGFR-like protein termed FGFR5 that lacks a KD, localises to the plasma membrane, is alternatively spliced and can bind FGFs and HS through its ectodomain, was first thought to act as a decoy receptor (Wiedemann and Trueb, 2000, Sleeman et al., 2001, Trueb et al., 2003). *FGFR5* knockout mice (but not those lacking only the intracellular domain) are not viable, so it is now thought that the ectodomain might mediate most of the activity of FGFR5, although kinase independent signalling roles have also been proposed based on the occurrence of a SH2 binding motif at its C-terminus (Trueb, 2011, Ornitz and Itoh, 2015, Kilkenny and Rocheleau, 2016).

It was already described in section 1.5.5 how the FGFR adaptor protein GRB2 might regulate signalling through ligand independent binding to the FGFR C-terminal tail. The FRS2 isoform FRS2α also constitutively (and ligand independently) binds to the intracellular FGFR juxtamembrane region via its PTB domain, from where it regulates downstream signalling (Xu et al., 1998a, Ong et al., 2000). Interestingly, GRB2 and FRS2 also interact with each other following ligand stimulation (Kouhara et al., 1997), and ternary complex formation between these two proteins and an E3 ubiquitin ligase in response to FGF stimulation leads to FGFR degradation and dampening of the transduced signal (Wong et al., 2002). Several other effectors also negatively regulate FGFR signalling by inhibiting the functions of signalling proteins downstream of FGFRs, as reviewed in (Ornitz and Itoh, 2015). It has even been proposed that extracellular ligand binding might be regulated in part by the intracellular FGFR KDs in an ‘inside-out’ like negative feedback mechanism (Uematsu et al., 2001). Regulation of FGFR signalling is clearly complex and multidimensional and is likely to be modulated in a context dependent manner in response to specific biological cues.

1.8 Deregulated FGFR signalling in pathology

Given the pleiotropic physiological roles of FGFR signalling, it is not surprising that deregulation of signalling leads to a range of diseases (Carter et al., 2015). First implications of deregulated FGFR signalling in pathophysiology came from the identification of gain-of-function germline mutations in FGFRs, which constitutively activated FGFRs in a ligand independent manner, and which were responsible for diverse skeletal dysplasias and craniosynostosis syndromes (Wilkie, 2005). In addition, deregulation of endocrine FGFR signalling was found to cause various metabolic defects, detailed in Chapter 4. However, it is the compelling evidence implicating aberrant FGFR signalling in cancer that has led to an explosion in the field of FGFR research in the last two decades. A study of over 1000 somatic mutations from over 200 different cancers highlighted the importance of FGF signalling in tumour biology (Greenman et al., 2007). Abnormal FGFR signalling can directly drive oncogenesis in a subset of several types of cancers by promoting cancer cell proliferation and survival and tumour angiogenesis (Turner and Grose, 2010, Dieci et al., 2013, Tanner and Grose, 2016). FGFRs are therefore considered important therapeutic targets in cancer (Brooks et al., 2012, Dienstmann et al., 2014, Touat et al., 2015).

1.8.1 Tumour diversity and types of FGFR aberrations

A recent study used next generation sequencing to comprehensively analyse nearly 5000 solid tumours from patients for aberrations in FGFRs, providing a representative overview of the landscape of FGFR deregulation in cancer (Helsten et al., 2016). This work showed that FGFR aberrations are found in 7% of all solid tumours, of which 5% and 2% represent gene amplifications and point mutations respectively. *FGFR1* was found to be the most frequently altered FGFR. In addition, FGFR aberrations were most often found in urothelial carcinomas, followed by breast, endometrial and ovarian carcinomas. In general, the most commonly occurring FGFR aberrations in cancer are gene amplifications, chromosomal fusions and somatic point mutations.

Several studies have shown that amplification and/or overexpression of *FGFR* genes appears to be the most common aberration found across several different cancers, with *FGFR1* being most frequently amplified (in breast and squamous cell lung cancers) followed by *FGFR2* (in endometrial cancer) and to a lesser degree *FGFR3* and *FGFR4* in a range of cancer types (Carter et al., 2015, Helsten et al., 2016). Chromosomal translocations involving *FGFR* genes are also commonly found in cancers, and in general these result in fusion proteins comprising the dimerisation domain of another protein fused to either the N- or C-terminus of FGFR, leading to constitutive and ligand independent FGFR dimerisation and activation (Turner and Grose, 2010, Gallo et al.,

2015). Oncogenic fusions were first found in bladder cancer and gliomas for FGFR3 and FGFR1 (Singh et al., 2012, Williams et al., 2013) and were subsequently also identified in several other cancers for FGFR1-3 (Wu et al., 2013). *FGFR2* is found fused to several different partners, whereas *FGFR3* is most commonly found fused to transforming acidic coiled-coil containing protein (*TACC3*) and this fusion produces one of the most important oncoproteins in bladder cancer (Gallo et al., 2015, di Martino et al., 2016, Babina and Turner, 2017). Somatic missense substitutions that result in point mutations are also observed very frequently in all FGFRs and in all types of cancer, but are most common in *FGFR3* and *FGFR2* and these occur most often in urothelial (bladder) and endometrial carcinomas respectively (Guagnano et al., 2012, Abbosh et al., 2015, Carter et al., 2015, Helsten et al., 2016). The most common point mutation in cancers found to date is the S249C mutation, in the extracellular D2-D3 linker of FGFR3, which accounts for more than 50% of all cancer-associated mutations in *FGFR3* (Tomlinson et al., 2007a, Babina and Turner, 2017). The breadth of FGFR point mutations in cancer and their activating nature and/or roles as oncogenic drivers in cancer will be discussed in detail in Chapter 3. Furthermore, the roles of aberrant endocrine FGFR signalling in cancer will be discussed in detail in Chapter 4.

1.8.2 Mechanisms of oncogenic FGFR signalling

There are common themes to the mechanisms by which FGFR aberrations induce oncogenic signalling (Figure 1.12). *FGFR* amplifications can lead to FGFR overexpression and accumulation at the plasma membrane, increasing activation of FGFRs and downstream signalling (Figure 1.12a). In addition, many FGFs have also been found overexpressed by tumour cells, or the adjacent tumour-associated stroma, leading to increased angiogenesis and epithelial-mesenchymal transition (EMT) through illicit autocrine or paracrine signalling (Figure 1.12b). Oncogenic fusion proteins, and certain point mutations, can force constitutive dimerisation of FGFRs leading to ligand independent hyperactive signalling (Figure 1.12c). Other types of point mutations can lead to allosteric or direct activation of the FGFR KDs leading to increased signalling in a ligand dependent or independent fashion, depending on the point mutation (Figure1.12d). Altered splicing of the autoinhibitory D1 domain of FGFRs and increased tumoural expression of this β splice isoform with a higher affinity for FGFs can lead to increased ligand dependent FGFR activation (Figure 1.12e). Similarly, a switch in expression from the FGFR b to the c splice isoform can alter the FGFR ligand specificity, allowing it to respond to a wider range of FGFs (for example those produced by the tumour cells) that it might not normally interact with (Figure 1.12f). It has even been found that an alternatively spliced FGFR2 variant lacking its C-terminus results in the loss of its ubiquitination site, impairing receptor internalisation and signal attenuation (Figure

1.12g). Finally, gene amplification, overexpression or mutation of intracellular FGFR effectors such as FRS2 or PLCγ can hyperactivate downstream signalling (Figure 1.12h).

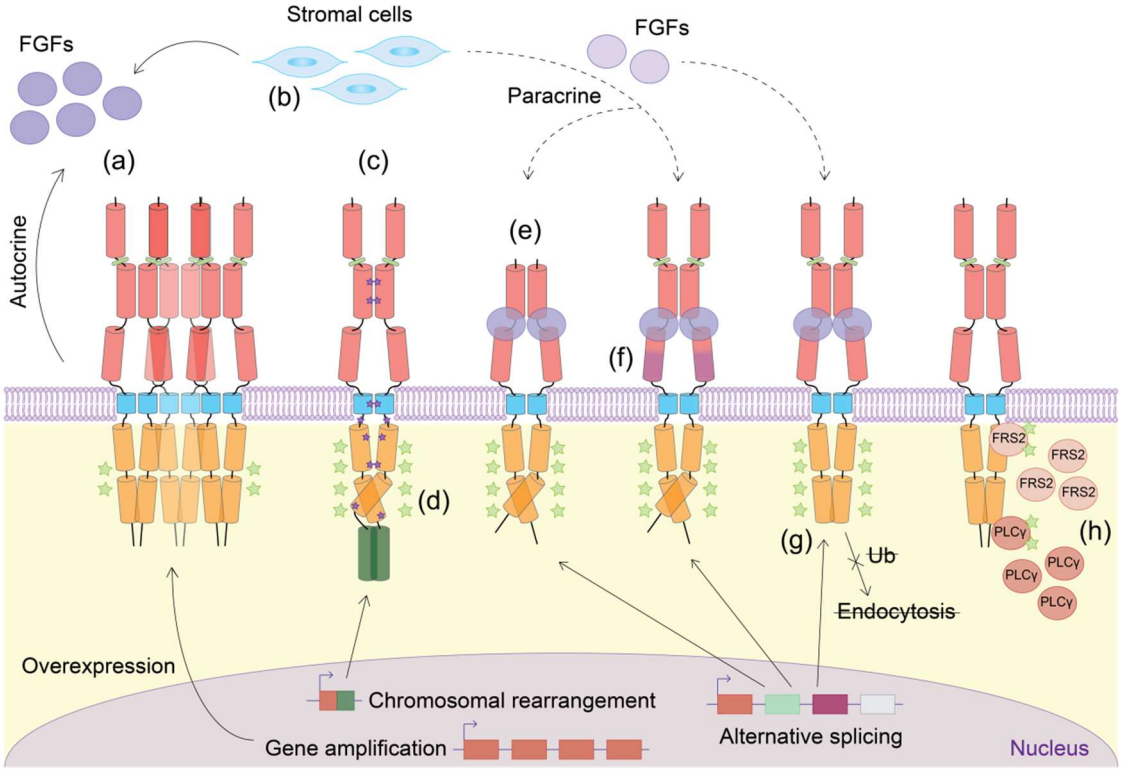


Figure 1.12 Mechanisms of oncogenic FGFR activation and signalling. There are several molecular mechanisms through which FGFR signalling can be aberrantly activated in cancer. **(a)** Amplification of the *FGFR* genes can lead to overexpression of receptors at the cell surface, leading to ligand-independent dimerisation by stochastic diffusion-limited activation (see also Figure 1.5a). **(b)** Amplification of *FGF* genes and/or overexpression of FGFs from surrounding (tumour-associated) stromal cells can also dimerise and activate FGFR signalling through illicit autocrine or paracrine routes. **(c)** Oncogenic fusion proteins (from chromosomal translocation or gene fusions, dark green) or activating point mutations in the extracellular, juxtamembrane or transmembrane domains (purple stars) can constitutively and ligand-independently dimerise and activate FGFRs. **(d)** Activating point mutations (purple stars) in the kinase domain (KD) can allosterically or directly activate the KD. **(e-g)** Alternative splicing of different parts of FGFRs can produce splice isoforms with altered ligand dependent or independent signalling properties, and relative expression of the more active splice variants is often found to be increased in tumours. **(e)** Alternative splicing of the D1 domain and/or the acid box in the D1-D2 linker relieves receptor autoinhibition and can increase affinity for FGFs. **(f)** Alternative splicing of the D3 domain can cause a switch from the alternative 'b' to the canonical 'c' FGFR splice isoforms. These can respond to a larger range of FGFs, such as those produced by tumour-associated stromal cells. **(g)** Alternative splicing of the C-terminal tail can remove crucial regulatory regions such that ubiquitination, recycling and/or degradation of activated FGFRs is impaired. **(h)** Gene amplification and/or overexpression of, or activating mutations in, downstream signalling effector or adaptor proteins can ligand-independently activate signalling.

1.8.3 Therapeutic targeting of aberrant FGFR signalling

The clear evidence that aberrant FGFR signalling is an important therapeutic target in several cancers has led to rapid progression in the field of anti-FGFR targeting. Despite these huge efforts, there are currently no approved therapies in the clinic based on FGFR targeting (Touat et al., 2015, Babina and Turner, 2017). A variety of therapeutic agents are in pre-clinical development, and encouragingly many are also at various stages of clinical trials for several cancer types with the aim of specific anti-FGFR targeting. These mostly comprise multi-kinase or FGFR-selective small molecule inhibitors which target the ATP binding pocket in the KD, but also include other therapeutic agents such as monoclonal antibodies (mAbs) and FGF 'ligand traps'. A list of anti-FGFR agents currently undergoing clinical trials is summarised in Table 1.4.

Apart from anti-FGFR therapies, the therapeutic potential of several FGFs has also been highlighted (Beenken and Mohammadi, 2009, Tanner and Grose, 2016). Anti-FGF therapeutics for various disorders are in pre-clinical development, while FGFs with beneficial effects on various aspects of physiology have been engineered to produce recombinant FGFs as biologic therapies. One such example is a recombinant form of FGF7 (palifermin) which is approved for the treatment of mucositis (Spielberger et al., 2004). In addition, the endocrine FGFs have been found to exert several beneficial metabolic effects and are rapidly being developed as therapeutics for diseases such as diabetes and obesity (detailed in Chapter 4). Approval of these candidate anti-FGFR and FGF treatments in the near future will help to address subsets of several diseases awaiting precision approaches to pathological targeting.

Table 1.4 Anti-FGFR therapeutics currently in clinical trials ($IC_{50(FGFR)}$ < 100 nM).

Inhibitor (company)	**Targets**	**Clinical development (cancer / tumour type)**
Multi-kinase (non-selective) small molecules		
Dovitinib (Novartis)	FGFR, KIT, VEGFR, PDGFR, RET	Phase II (gastric, urothelial) Phase III (metastatic renal)
Ponatinib (ARIAD Pharmaceuticals)	FGFR, ABL, VEGFR, PDGFR, SRC	Phase II (biliary, advanced solid)
Lenvatinib (Eisai)	FGFR, VEGFR, PDGFR	Phase III (thyroid)
Lucitanib (Clovis Oncology)	FGFR, VEGFR, PDGFR	Phase I (advanced solid) Phase II (metastatic breast / lung)
Nintedanib (Boehringer Ingelheim)	FGFR, VEGFR, PDGFR	Phase III (NSCLC, ovarian)
Pan-FGFR (FGFR-selective) small molecules		
AZD4547 (AstraZeneca)	FGFR1-3	Phase II (SCLC, NSCLC, metastatic NSCLC)
NVP-BGJ398 (Novartis)	FGFR1-3	Phase I (SCLC, urothelial) Phase II (solid, haematological)
JNJ-42756493 (Janssen)	FGFR1-4	Phase I (*FGF19* amplified HCC*) Phase II (urothelial, Asian solid)
LY2874455 (Eli Lilly)	FGFR1-4	Phase I (advanced solid)
ARQ-087 (ArQule)	FGFR1-3	Phase III (bile ductal and/or HCC)
Debio-1347 (Debiopharm)	FGFR1-3	Phase I (advanced solid)
TAS120 (Taiho Oncology)	FGFR1-4	Phase I (advanced solid) Phase II (myeloma)
E7090 (Eisai)	FGFR1-3	Phase I (advanced solid)
ASP5878 (Astellas Pharma)	FGFR1-4	Phase I (advanced solid)
Anti-FGFR antibodies		
R3Mab (Genentech)	FGFR3	Phase I (advanced solid)
FPA144 (Five Prime Therapeutics)	FGFR2b	Phase I (advanced solid)
BAY1179470 (Bayer)	FGFR2	Phase I (advanced solid)
FGF ligand traps (soluble FGFR ectodomains)		
FP-1039 (Five Prime Therapeutics)	FGF2 and other FGFs	Phase I (advanced solid)

Chapter 2: Materials and Methods

2.1 General molecular biology methods

2.1.1 Production of competent *E. coli* cells

Escherichia coli (*E. coli*) cells (specific strains detailed in respective sections) were streaked from existing lab competent cell stocks onto lysogeny broth (LB) agar plates (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl and 15 g/L agar) containing the appropriate antibiotics and incubated overnight at 37°C. A single colony was used to inoculate a 25 mL LB medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl) starter culture for 7 hours at 37°C. The starter culture was used to inoculate 250 mL SOB medium (20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 2.5 mM KCl and 10 mM $MgSO_4$) supplemented with 10 mM $MgCl_2$ at different cell densities overnight at 18°C. When cell density reached an optical density of 0.55 at 600 nm (OD_{600}) the culture was transferred to ice for 10 min. Cells were harvested by centrifugation (4,000 x *g*, 10 minutes and 4°C) and then resuspended in 80 mL ice-cold Inoue transformation buffer (ITB: 10 mM PIPES-KOH pH 6.7, 250 mM KCl, 15 mM $CaCl_2$ and 55 mM $MnCl_2$) by swirling gently for 10 minutes at 4°C. Cells were harvested by centrifugation (4,000 x *g*, 10 minutes and 4°C) and resuspended in 20 mL ice-cold ITB again by swirling gently for 10 minutes at 4°C. DMSO was added to the cell suspension to 7% v/v. Cell suspension was aliquoted (50 µL and 250 µL), snap frozen in liquid nitrogen and stored at -80°C.

2.1.2 Bacterial transformation

DNA (50-300 ng in 0.5-2 µL) was mixed with 50 µL competent *E. coli* cells (specific strains detailed in respective sections) for 30 minutes on ice. The mixture was subjected to a 'heat-shock' (42°C for 45 seconds) before being incubated for a further 2 minutes on ice. Transformed cells were then shaken for 1 hour at 37°C after the addition of 250 µL SOC medium (SOB medium supplemented with 20 mM $MgCl_2$ and 10 mM D-glucose). Cell suspension was then spread on agar plates comprising appropriate growth medium and containing the appropriate antibiotics (10 mM D-glucose was also added to the plates when necessary to strictly repress basal expression of proteins toxic to the *E. coli*) and incubated overnight at 37°C.

2.1.3 Restriction cloning

Restriction digests (to prepare vector or insert for cloning) were performed with 1 µg template plasmid DNA, 1 x reaction buffer (NEB), 0.1 mg/mL BSA and 1 µL of each appropriate restriction enzyme (NEB) per 25 µL digest. Reactions were heat-inactivated (65°C for 5 min) and run on 0.9% w/v agarose gels pre-stained with GelRed™ (Biotium).

Digested DNA was visualised under a UV-lamp and bands at the correct molecular weights were cut and extracted from the gel using the QIAquick Gel Extraction kit (Qiagen) as per the manufacturer's instructions. Vector and insert were then mixed in a 3:1 molar ratio for ligation with T4 ligase (NEB) and the reaction was incubated overnight at 16°C. Competent *E. coli* XL10 Gold cells were then transformed with 1 µL of ligated product and colonies were sequenced to screen for successful cloning.

2.1.4 Ligation-independent cloning (LIC)

Linearised DNA vectors and inserts were generated using 50 ng template plasmid DNA, appropriate primers and PfuUltra™ DNA polymerase (Stratagene) as per the manufacturer's instructions, with reactions performed on a thermal cycler (G storm). When required, polymerase chain reactions (PCRs) were performed using temperature gradients. Total PCR product was run on 0.9% w/v agarose gels pre-stained with GelRed™ dye (Biotium). Linearised DNA was visualised under a UV-lamp and bands at the correct molecular weights were cut and extracted from the gel using the QIAquick Gel Extraction kit (Qiagen) as per the manufacturer's instructions. Vector and insert were then mixed in variable (optimised) amounts and incubated with In-Fusion® HD buffer/enzyme mix (Clontech) for 15 minutes at 37°C followed by 15 minutes at 50°C. Competent *E. coli* XL10 Gold cells were then transformed with 1 µL LIC product and colonies were sequenced to screen for successful cloning.

2.1.5 Site-directed mutagenesis (SDM)

SDM reactions were performed on a thermal cycler (G storm) using 5 ng template plasmid DNA, appropriate primers and PfuUltra™ DNA polymerase (Stratagene) as per the manufacturer's instructions. When required, PCR was performed using appropriate temperature gradients. Total PCR product was subjected to digestion with DpnI (NEB) for 1 hour at 37°C to eliminate remaining template DNA. Competent *E. coli* XL10 Gold cells were then transformed with 2 µL digested PCR product and colonies were sequenced to screen for successful incorporation of mutations.

2.1.6 Large scale plasmid purification

A 3 mL starter culture (inoculated with *E. coli* XL10 Gold cells transformed with the relevant plasmid) was used to inoculate cultures of 500 mL LB medium or terrific broth (TB) medium (12 g/L tryptone, 24 g/L yeast extract, 0.4% v/v glycerol, 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) shaken overnight at 37°C. Cells were harvested by centrifugation (4,500 x *g*, 10 minutes and 4°C) and resuspended in 37 mL alkaline lysis buffer I (50 mM D-glucose, 25 mM Tris-HCl pH 8.0 and 10 mM EDTA). Cells were lysed with 50 mL alkaline lysis buffer II (0.2 M NaOH and 1% w/v SDS) and genomic DNA precipitated

with 37 mL cold alkaline lysis buffer III (3 M potassium acetate and 11.5% v/v glacial acetic acid) followed by centrifugation (4,500 x *g*, 10 minutes and 4°C). Supernatant was filtered through 4 layers of cheesecloth and nucleic acid precipitated with 0.7 volumes propan-2-ol followed by centrifugation (4,500 x *g*, 15 minutes and 4°C). The pellet was washed in 10 mL 70% ethanol, dried and re-dissolved in 2 mL TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0). RNA was precipitated with 2 mL 5 M LiCl followed by centrifugation (4,500 x *g*, 10 minutes and 4°C). All remaining nucleic acid was precipitated from the supernatant with 0.7 volumes propan-2-ol followed by centrifugation (4,500 x *g*, 10 minutes and 4°C). Pelleted nucleic acids were re-dissolved in 0.5 mL TE buffer and treated with 5 µL of 10 mg/mL RNase A (Qiagen) for 15 minutes at 37°C. Undigested nucleic acids were precipitated with 0.5 mL 13% w/v polyethylene glycol 8,000 and 1.5 M NaCl for 30 minutes on ice followed by centrifugation (4,500 x *g*, 5 minutes and room temperature). Dried pellet was re-dissolved in 200 µL TE buffer and added to 200 µL phenol/chloroform (Sigma), vortexed for 30 seconds and centrifuged (20,000 x *g*, 5 minutes and room temperature). The upper phase was transferred to 0.1 volumes 3 M sodium acetate and 2.5 volumes 100% ethanol added. Following a final centrifugation (20,000 x *g*, 2 minutes and 4°C) the pellet was washed with 0.5 mL 70% ethanol, dried completely and re-dissolved in 500 µL TE buffer. DNA concentration was quantified on a Nanodrop 2000c and plasmid preparations were adjusted to 1 µg/µL with additional TE buffer.

Alternatively, 150 mL cultures were similarly grown up, cells harvested and plasmids purified using the ZymoPURE™ Plasmid Maxiprep kit (Zymo Research) as per the manufacturer's instructions.

2.2 Construction of plasmid libraries

2.2.1 FGFR3 and variants

Native human FGFR3 (455-768) kinase domain (KD) cDNA was obtained from a lab stock in the pOPINS vector (OPPF) which contains an N-terminal Ulp1 protease cleavable 6His-SUMO tag. Solubilising mutations C482A and C582S (Mohammadi et al., 1996b) were introduced by sequential SDM by Tom Bunney to produce the '2C' mutations C482A and C582S and thus FGFR3-2C (455-768). Native human FGFR3b full length cDNA was cloned in the pFB vector (originally pFB-neomycin resistant, Stratagene) by the Knowles lab, University of Leeds, modified to be hygromycin B resistant (Tomlinson et al., 2005). Missense substitutions were then introduced into both pOPINS (kanamycin-resistant) and pFB (ampicillin-resistant) constructs by SDM using

the primers in Table 2.1. Some of these mutant plasmids were generated by Sarah Saad and Tom Bunney (pOPINS) and Yulia Panina (pFB) respectively.

A codon-optimised human FGFR3-2C (455-768) gene harbouring an additional K650E mutation was synthesised in the pJ821 vector (DNA2.0) with an N-terminal cleavable 6His-SUMO tag to produce a similar construct to the pOPINS construct. The K650E mutation was first mutated back to lysine to recreate FGFR3-2C using the primers in Table 2.2a. Some of the mutations from Table 2.1, and some mutations selected for analysis at a later date, were then introduced into the codon-optimised pJ821 FGFR3-2C (455-768) backbone (kanamycin-resistant) by SDM using the primers in Table 2.2b.

For pOPINS and pJ821 constructs, individual colonies were picked and used to inoculate small volume overnight cultures at 37°C, from which glycerol stocks (containing 25% v/v glycerol mixed with cell suspension) were prepared and plasmids were purified using the QIAprep Spin Miniprep kit (Qiagen) as per the manufacturer's instructions. Purified plasmids were pre-sequenced using the BigDye3.1 system (ABI Prism) and the DyeEx 2.0 Spin kit (Qiagen) as per the manufacturer's instructions, and sequenced (Source BioScience) to confirm perfect incorporation of mutations. For pFB constructs, individual colonies were picked and subjected to colony PCR using Taq DNA polymerase (Qiagen) and PCR product was analysed by 1% w/v agarose gel electrophoresis. Colonies with detectable PCR product were used to inoculate small volume overnight cultures at 37°C, and glycerol stocks and plasmids were prepared as detailed above. Purified plasmids were then pre-sequenced and deep sequenced as described above to retain constructs with perfectly incorporated mutations and discard false positives.

Table 2.1 Site-directed mutagenesis primers used to introduce missense substitutions into the native pOPINS FGFR3-2C (455-768) and pFB FGFR3b (full length) backbone constructs. Residue numbering is as per the canonical FGFR3c splice isoform.

FGFR3 mutation		**Primer**	
E466K	GAG → **A**AG	Forward	CCGACCCCAAATGG**A**AGCTGTCTCGGGCCC
		Reverse	GGGCCCGAGACAGCT**T**CCATTTGGGGTCGG
A500T	GCC → **A**CC	Forward	GACAAGGACCGGGCC**A**CCAAGCCTGTCACC
		Reverse	GGTGACAGGCTTGG**T**GGCCCGGTCCTTGTC
N540K	AAC → AA**G**	Forward	CAAAAACATCATCAA**G**CTGCTGGGCGCCTGC
		Reverse	GCAGGCGCCCAGCAG**C**TTGATGATGTTTTTG
N540S	AAC → A**G**C	Forward	CAAAAACATCATCA**G**CCTGCTGGGCGCCTG
		Reverse	CAGGCGCCCAGCAGG**C**TGATGATGTTTTTG
P572A	CCC → **G**CC	Forward	GCGGGCGCGGCGG**G**CCCCGGGCCTGGACTAC
		Reverse	GTAGTCCAGGCCCGGGG**C**CCGCCGCGCCCGC
C582F	TGC → T**T**C	Forward	CTCCTTCGACACCT**T**CAAGCCGCCCGAGGAG
		Reverse	CTCCTCGGGCGGCTTG**A**AGGTGTCGAAGGAG
D617G	GAC → G**G**C	Forward	GTGCATCCACAGGG**G**CCTGGCTGCCCGCAATG
		Reverse	CATTGCGGGCAGCCAGG**C**CCCTGTGGATGCAC
E627D	GAG → GA**C**	Forward	GTGCTGGTGACCGA**C**GACAACGTGATGAAG
		Reverse	CTTCATCACGTTGTC**G**TCGGTCACCAGCAC
V630M	GTG → **A**TG	Forward	GACCGAGGACAAC**A**TGATGAAGATCGCAGAC
		Reverse	GTCTGCGATCTTCATCA**T**GTTGTCCTCGGTC
D641G	GAC → G**G**C	Forward	GGCTGGCCCGGG**G**CGTGCACAACCTC
		Reverse	GAGGTTGTGCACG**C**CCCGGGCCAGCC
D641N	GAC → **A**AC	Forward	CGGGCTGGCCCGG**A**ACGTGCACAACCTCGAC
		Reverse	GTCGAGGTTGTGCACGT**T**CCGGGCCAGCCCG
H643D	CAC → **G**AC	Forward	GGCCCGGGACGTG**G**ACAACCTCGACTACTAC
		Reverse	GTAGTAGTCGAGGTTGT**C**CACGTCCCGGGCC
D646Y	GAC → **T**AC	Forward	CGTGCACAACCTC**T**ACTACTACAAGAAGAC
		Reverse	GTCTTCTTGTAGTAGT**A**GAGGTTGTGCACG
K650E	AAG → **G**AG	Forward	CGACTACTACAAG**G**AGACAACCAACGGCCG
		Reverse	CGGCCGTTGGTTGTCT**C**CTTGTAGTAGTCG
K650M	AAG → A**T**G	Forward	GACTACTACAAGA**T**GACAACCAACGGCCGG
		Reverse	CCGGCCGTTGGTTGTC**A**TCTTGTAGTAGTC
K650N	AAG → AA**C**	Forward	GACTACTACAAGAA**C**ACAACCAACGGCCGG
		Reverse	CCGGCCGTTGGTTGT**G**TTCTTGTAGTAGTC
N653H	AAC → **C**AC	Forward	CAAGAAGACAACC**C**ACGGCCGGCTGCCCGTG
		Reverse	CACGGGCAGCCGGCCGT**G**GGTTGTCTTCTTG
R669G	CGA → **G**GA	Forward	GAGGCCTTGTTTGAC**G**GAGTCTACACTCACC
		Reverse	GGTGAGTGTAGACTC**C**GTCAAACAAGGCCTC
R669Q	CGA → C**A**A	Forward	GAGGCCTTGTTTGACC**A**AGTCTACACTCACCAG
		Reverse	CTGGTGAGTGTAGACT**T**GGTCAAACAAGGCCTC
V677I	GTC → **A**TC	Forward	CTCACCAGAGTGAC**A**TCTGGTCCTTTGGGG
		Reverse	CCCCAAAGGACCAGA**T**GTCACTCTGGTGAG
E686K	GAG → **A**AG	Forward	GGTCCTGCTCTGG**A**AGATCTTCACGCTGGG
		Reverse	CCCAGCGTGAAGATCT**T**CCAGAGCAGGACC
G697C	GGC → **T**GC	Forward	CTCCCCGTACCCC**T**GCATCCCTGTGGAGGAG
		Reverse	CTCCTCCACAGGGATGC**A**GGGGTACGGGGAG

Table 2.2 Site-directed mutagenesis primers used to (**a)** reintroduce a native lysine at position 650 of the codon-optimised pJ821 FGFR3-2C (455-768) harbouring a K650E mutation, and (**b)** introduce missense substitutions into the pJ821 FGFR3-2C (455-768) backbone. Residue numbering is as per the canonical FGFR3c splice isoform.

(a)

FGFR3 mutation		**Primer**	
E650K	GAG → **A**AG	Forward	CCTGGACTACTATAAA**A**AGACTACTAATGGCCGC
		Reverse	GCGGCCATTAGTAGTCT**T**TTTATAGTAGTCCAGG

(b)

FGFR3 mutation		**Primer**	
I538F	ATC → **T**TC	Forward	CGGCAAGCACAAGAAC**T**TCATCAATCTGCTGGGTGC
		Reverse	GCACCCAGCAGATTGATGA**A**GTTCTTGTGCTTGCCG
I538V	ATC → **G**TC	Forward	CGGCAAGCACAAGAAC**G**TCATCAATCTGCTGGGTGC
		Reverse	GCACCCAGCAGATTGATGA**C**GTTCTTGTGCTTGCCG
N540K	AAT → **A**AA	Forward	CAAGAACATCATCAA**A**CTGCTGGGTGCTTGTACC
		Reverse	GGTACAAGCACCCAGCAG**T**TTGATGATGTTCTTG
G637W	GGC → **T**G**G**	Forward	GAAAATTGCGGACTTC**T**G**G**TTGGCACGCGATGTTC
		Reverse	GAACATCGCGTGCCAA**C**C**A**GAAGTCCGCAATTTTC
D641G	GAT → G**G**T	Forward	CGGCTTGGCACGCG**G**TGTTCACAACC
		Reverse	GGTTGTGAACA**C**CGCGTGCCAAGCCG
R669G	CGT → **G**GT	Forward	GGAAGCATTGTTTGAT**G**GTGTGTACACCCACC
		Reverse	GGTGGGTGTACACAC**C**ATCAAACAATGCTTCC
R669Q	CGT → C**AG**	Forward	GGAAGCATTGTTTGATC**AG**GTGTACACCCACC
		Reverse	GGTGGGTGTACAC**CT**GATCAAACAATGCTTCC
E686K	GAG → **A**AG	Forward	GGTGTCCTGTTGTGG**A**AGATCTTTACGCTG
		Reverse	CAGCGTAAAGATCT**T**CCACAACAGGACACC

2.2.2 β-Klotho, FGFR1c, FGF19 and FGF21

Codon-optimised constructs of human and mouse βKlotho, human FGFR1c and human FGF19 and FGF21 were custom designed in the bacterial vector pAZET09 (AstraZeneca) or mammalian vectors pcDNA3.1 (Invitrogen) and pEBNAZ (MedImmune) and cloned by GeneArt™ (Thermo Fisher Scientific). All constructs comprised a C-terminal protease cleavable 2StrepII-10His tag. A full length native human FGFR1c cDNA construct in the pcDNA3.1 mammalian vector with a C-terminal FLAG tag was provided by AstraZeneca. Additionally a full length native mouse βKlotho cDNA construct in a pCMV mammalian vector with a C-terminal fusion to eGFP was purchased (Addgene). Restriction enzyme cloning, LIC and SDM were used to modify these constructs (e.g. introduce or remove fusion tags or sub-clone constructs into different expression vectors). Details of all constructs prepared for use in this study are in Tables 2.3 and 2.4.

Table 2.3 Expression constructs prepared for expression testing in *E. coli* (specific strains detailed in respective sections) and for recombinant protein production in this work. All constructs have a C-terminal PreScission™ protease-cleavable 2StrepII-10His tag. The gene from which the N-terminal secretion signal peptide was derived, where applicable, is rendered in italics. Refer to text for full details.

Protein	Species	Construct	Residues	Expression vector	Source
FGF19	Human	FL	23-216	pAZET09	GeneArt™
FGF21	Human	FL	29-209	pAZET09	GeneArt™
βKlotho	Human	*HSA*-ΔKL1	513-1044	pAZET09	GeneArt™
βKlotho	Human	*pelB*-ΔKL1	513-1044	pAZET09	GeneArt™
βKlotho	Human	*pelB*-KL1	53-514	pAZET09	GeneArt™
βKlotho	Human	KL1	53-514	pAZET09	GeneArt™
βKlotho	Human	*pelB*-KL2	513-972	pAZET09	GeneArt™
βKlotho	Human	KL2	513-972	pAZET09	GeneArt™
βKlotho	Human	*HSA*-FL	53-1044	pAZET09	GeneArt™
βKlotho	Mouse	*HSA*-FL	53-1043	pAZET09	GeneArt™
βKlotho	Human	*pelB*-FL	53-1044	pAZET09	GeneArt™
βKlotho	Human	*HSA*-KL1-KL2	53-995	pAZET09	GeneArt™
βKlotho	Human	*pelB*-KL1-KL2	53-995	pAZET09	GeneArt™
βKlotho	Human	KL1-KL2	53-995	pAZET09	GeneArt™
βKlotho	Mouse	*pelB*-FL	53-1043	pAZET09	GeneArt™
βKlotho	Human	*HSA*-FL	53-1044	pTriEx™4	Subcloned
βKlotho	Human	FL	53-1044	pTriEx™4	Subcloned
βKlotho	Human	*HSA*-KL1-KL2	53-995	pTriEx™4	Subcloned
βKlotho	Human	KL1-KL2	53-995	pTriEx™4	Subcloned
FGFR1c	Human	D1-D3	22-374	pAZET09	Subcloned
FGFR1c	Human	D2-D3	142-365	pAZET09	Subcloned

2.3 General *E. coli* recombinant protein expression and purification

2.3.1 Recombinant protein expression

Cultures of 500 mL TB medium containing the appropriate antibiotics were inoculated with transformed *E. coli* from glycerol stocks (specific strains detailed in respective sections) or multiple colonies from freshly transformed *E. coli* and shaken at 37°C at 200-250 revolutions per minute (rpm). When cell density reached an OD_{600} of 1.0, 0.4 mM IPTG (or 1 mM L-rhamnose in the case of pJ821 constructs) was added to induce expression and cells were incubated overnight at 15°C, 20°C or 37°C as appropriate. Cells were harvested by centrifugation (4,500 x *g*, 20 minutes and 4°C) and pellets were frozen at -20°C for at least 1 hour before lysis.

Table 2.4 Expression constructs prepared for expression testing in mammalian cell lines and for recombinant protein production in this work. The gene from which the N-terminal secretion signal peptide was derived, where applicable, is rendered in italics. The alias provided for some constructs is used instead of the construct name in some Figures in Chapter 4. FL – full length, HSA – human serum albumin.

Protein	Species	Construct [Alias]	Residues	Expression vector	C-terminal fusion tag	Source
βKlotho	Human	*βKL*-FL [41]	1-1044	pcDNA3.1	2StrepII-10His	GeneArt™
βKlotho	Human	*βKL*-KL1-KL2 [42]	1-995	pcDNA3.1	2StrepII-10His	GeneArt™
βKlotho	Human	*CD33*-KL1-KL2 [43]	53-995	pcDNA3.1	2StrepII-10His	GeneArt™
βKlotho	Human	*CD33*-KL1-KL2 [K3]	53-995	pEBNAZ	2StrepII-10His	GeneArt™
βKlotho	Mouse	FL [mβe]	1-1043	pCMV	*eGFP	Addgene
βKlotho	Mouse	FL [mβTe6H]	1-1043	pCMV	eGFP-6His	Engineered
βKlotho	Human	*HSA*-FL	53-1044	pTriEx™4	2StrepII-10His	Subcloned
βKlotho	Human	FL	53-1044	pTriEx™4	2StrepII-10His	Subcloned
βKlotho	Human	*HSA*-KL1-KL2	53-995	pTriEx™4	2StrepII-10His	Subcloned
βKlotho	Human	KL1-KL2	53-995	pTriEx™4	2StrepII-10His	Subcloned
FGFR1c	Human	FL	1-822	pcDNA3.1	*FLAG	AstraZeneca
FGFR1c	Human	D1-D3	22-374	pcDNA3.1	*FLAG	Engineered
FGFR1c	Human	D2-D3	142-365	pcDNA3.1	*FLAG	Engineered
FGFR1c	Human	*CD33*-D1-D3 [45]	22-374	pcDNA3.1	2StrepII-10His	GeneArt™
FGFR1c	Human	*CD33*-D1-D3 [F1]	22-374	pEBNAZ	2StrepII-10His	GeneArt™
FGFR1c	Human	*CD33*-D2-D3 [46]	142-365	pcDNA3.1	2StrepII-10His	GeneArt™

* no TEV protease cleavage site (rest have TEV protease cleavage site)

2.3.2 Chemical lysis

Harvested frozen *E. coli* cells (specific strains detailed in respective sections) were resuspended by shaking for 1 hour at 4°C in lysis buffer (25 mM Tris-HCl pH 8.0, 250 mM NaCl, 40 mM imidazole, 10 mM benzamidine hydrochloride, 1 mM $MgCl_2$, 0.5 mM $CaCl_2$ and 1 mg/mL lysozyme) and then lysed by the addition of 2% v/v cold Triton X-100 and 15 mg/L DNase I for another 1 hour at 4°C. Lysate was clarified by centrifugation (20,000 x *g*, 1 hour and 4°C) to separate soluble protein (supernatant) from insoluble protein in inclusion bodies (pellet).

2.3.3 Small-scale expression tests in *E. coli*

Overnight 1 mL cultures were inoculated from glycerol stocks transformed with the relevant expression constructs in the relevant *E. coli* expression strains and were shaken at 37°C in LB medium containing the relevant antibiotic. Overnight cultures were used to inoculate 50 mL TB medium cultures containing the relevant antibiotic (normally 100 mg/L kanamycin or 50 mg/L ampicillin) and cultures were shaken at 37°C. Alternatively, no overnight starter culture was used and 2-5 mL TB medium cultures containing the relevant antibiotics were directly inoculated from glycerol stocks and shaken at 37°C. Following induction of expression with 0.5 mM IPTG, cells were shaken overnight at the required temperature (15, 20, 25 or 37°C) or for 3-4 hours at 37°C. Cells were then harvested and frozen. Harvested frozen cell pellets were then thawed on ice, lysed by chemical lysis and separated into soluble (S) and insoluble (IS) fractions. S and IS fractions were then analysed for protein expression by SDS-PAGE analysis and anti-5His immunoblotting.

2.4 SDS-PAGE

Protein samples were prepared by mixing 1 volume part 4 x SDS-PAGE sample buffer (0.2 M Tris-HCl pH 6.8, 8% w/v SDS, 0.05% w/v bromophenol blue, 40% v/v glycerol and 5% 2-mercaptoethanol) with 3 volume parts protein sample. Samples were then boiled for 5-10 minutes at 95°C before the denatured samples were loaded onto 7.5%, 10%, 12% or 15% SDS-PAGE gels (produced in-house) depending on the application. Gels were run using Mini-PROTEAN® electrophoresis cells (Bio-Rad) with SDS-PAGE running buffer (25 mM Tris, 192 mM glycine and 0.1% w/v SDS).

2.5 Immunoblotting

Protein samples were resolved by SDS-PAGE. Gels were then soaked in transfer buffer (25 mM Tris, 192 mM glycine and 20% v/v methanol) for 2 minutes. Proteins were transferred from gel to PVDF or nitrocellulose membranes using either Trans-Blot® SD Semi-Dry Transfer Cell or the Mini-PROTEAN® Trans-Blot Cells (Bio-Rad). Membranes were blocked in 5% w/v non-fat dried milk (NFDM) or bovine serum albumin (BSA) respectively in TBS-T buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.1% v/v Tween® 20) depending on the antibody to be used. Membranes were incubated with primary antibodies (details in Table 2.5) in the corresponding blocking solution overnight at 4°C. All primary antibodies were used with HRP-conjugated secondary antibodies and ECL™ Prime (GE Healthcare) and imaged using the Odyssey® Fc system (Li-Cor

Biosciences) or Super RX Fuji medical x-ray film (Kodak). Densitometric analyses were performed using Image Studio™ Lite (Li-Cor Biosciences) or ImageJ software.

Table 2.5 Primary antibodies and experimental conditions used for immunoblotting. mAb (monoclonal antibody), pAb (polyclonal antibody), BSA (bovine serum albumin), NFDM (non-fat dried milk), CST (cell signaling technologies).

Primary antibody (clone)	Origin	Working Dilution	Blocking Conditions	Source
5His	Mouse mAb	1:5000	5% BSA	Qiagen, #34660
β-actin	Rabbit pAb	1:2000	5% NFDM	Abcam, #75186
FGFR3 (B9)	Mouse mAb	1:1000	5% NFDM	Santa Cruz, #sc-13121
ERK1/2	Rabbit pAb	1:2000	5% NFDM	Millipore, #06-182
pERK1/2 (T202/Y204, T185/Y187)	Rabbit mAb	1:2000	5% BSA	CST, #4370
AKT1/2/3	Rabbit pAb	1:1000	5% NFDM	CST, #9272
pAKT1/2/3 (S473, S474, S472)	Rabbit pAb	1:1000	5% BSA	CST, #9271
PLCγ1	Rabbit pAb	1:1000	5% NFDM	CST, #2822
FGFR1	Mouse / Rabbit pAb*	1:1000*	5% NFDM	Generated in-house
KLB (βKlotho human, mouse)	Rabbit pAb	1:1000	5% NFDM	OriGene, #TA337557
mβKL (βKlotho mouse)	Goat pAb	1:5000	5% NFDM	Bio-Techne, #AF2619
GFP	Rabbit pAb	1:1000	5% NFDM	Santa Cruz, sc-8334
GFP (B2)	Mouse mAb	1:1000	5% NFDM	Santa Cruz, #sc-9996

*created by mixing 3 x mAbs, with each mAb at 1:1000.

2.6 Expression and purification of recombinant proteins from *E. coli*

2.6.1 Human FGFR3-2C (455-768) and variant KD proteins

Competent C41 (DE3) λ37 *E. coli* cells (pre-transformed with a spectinomycin-resistant pCDFDuet-1™ accessory plasmid [Novagen] containing genes for λ-phosphatase and CDC37 co-chaperone) were transformed with 6His-SUMO-FGFR3-2C (455-768) and

mutant constructs in pOPINS or pJ821 vectors. LB agar plates used for transformations of pOPINS constructs contained 10 mM D-glucose to strictly repress basal recombinant protein expression during growth on the solid medium, alongside the relevant antibiotics. Multiple colonies were used to inoculate cultures of 500 mL TB medium containing kanamycin (100 mg/L) and spectinomycin (50 mg/L) and cultures were shaken at 37 °C. Following induction of expression with 0.4 mM IPTG (pOPINS constructs) or 1 mM L-rhamnose (pJ821 constructs) and/or an additional 0.1 mM IPTG (for co-expression of accessory plasmid pCDFDuet-1™ with pJ821 constructs) cells were shaken overnight at 15°C. Cells were then harvested and frozen. Harvested frozen cell pellets were thawed on ice and lysed by chemical lysis. Following lysate clarification, supernatant was loaded onto a HisTrap FF Crude 5 mL column (GE Healthcare) pre-equilibrated in His buffer A (25 mM Tris-HCl pH 8.0, 500 mM NaCl and 40 mM imidazole). Protein was eluted in an imidazole gradient with 0-100% His buffer B (25 mM Tris-HCl pH 8.0, 500 mM NaCl and 500 mM imidazole) and the proteins eluted at 150-200 mM imidazole. Protein eluate was dialysed overnight in dialysis buffer (25 mM Tris-HCl pH 8.0 and 1 mM TCEP) at 4°C in 10 kDa molecular weight cut-off (MWCO) snakeskin dialysis tubing (Thermo Fisher Scientific) in the presence of 50 µg/mL Ulp1 protease (produced in-house) to cleave off the N-terminal 6His-SUMO tag. Cleaved protein was loaded onto a HiTrap Chelating HP 5 mL column (GE Healthcare) pre-equilibrated in dialysis buffer and eluted cleaved and tag-free in this buffer. Cleaved eluted protein was then subjected to anion exchange on a HiTrap Q HP 5 mL column (GE Healthcare) pre-equilibrated in Q buffer A (25 mM Tris-HCl, 20 mM NaCl and 1 mM TCEP). Protein was eluted in a NaCl gradient with 0-50% Q buffer B (25 mM Tris-HCl, 1 M NaCl and 1 mM TCEP) and protein eluted at 200-250 mM NaCl. Fractions were analysed by non-reducing 12% SDS-PAGE and relevant peak fractions were pooled and concentrated to 3 mL using Vivaspin-20 10 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein was injected onto a HiLoad 16/60 Superdex 75 prep grade column (GE Healthcare) for size exclusion chromatography (SEC), pre-equilibrated in SEC buffer (25 mM Tris-HCl, 150 mM NaCl and 1 mM TCEP). Peak fractions were analysed by non-reducing 12% SDS-PAGE and then pooled and concentrated to 100-500 µL, 2-10 mg/mL and over 95% purity in most cases (purity estimated by SDS-PAGE analysis) using fresh Vivaspin-6 10 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein concentrations were measured using the Nanodrop 2000c (Thermo Fisher Scientific). Protein was aliquoted in 50-100 µL aliquots, snap frozen in liquid nitrogen and stored at -80°C.

2.6.2 Human FGF1 (16-155) protein

Expression and purification of pOPINS-6His-SUMO-FGF1 (16-155) (DNA construct obtained from Tom Bunney) was performed very similarly to methods used to express

and purify FGFR3-2C (455-768) and mutant KD proteins, with a few modifications. C41 (DE3) *E. coli* cells were used, therefore TB medium was supplemented only with kanamycin (100 mg/L). For dialysis steps, 3.5 kDa MWCO snakeskin dialysis tubing (Thermo Fisher Scientific) was used. For SEC, a non-reducing (NR) SEC buffer (25 mM Tris-HCl and 137 mM NaCl) was used and peak fractions were analysed by reducing 15% SDS-PAGE. Peak fractions were pooled and concentrated to 1500 µL, 4.3 mg/mL and over 95% purity (purity estimated by SDS-PAGE analysis) using fresh Vivaspin-20 3.5 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech).

2.6.3 Human βKlotho KL1 (53-514) and KL2 (513-972) proteins

Competent C41 (DE3) *E. coli* cells were transformed with a KL1 (53-514) or KL2 (513-972) construct in the pAZET09 vector. An overnight 10 mL culture inoculated from a glycerol stock was used to inoculate 500 mL TB medium cultures containing kanamycin (100 mg/L) and cultures were shaken at 37°C. Following induction of expression with 0.5 mM IPTG cells were shaken overnight at 20°C. Cells were then harvested and frozen. Harvested frozen cell pellets were thawed on ice and lysed by chemical lysis. Following lysate clarification, supernatant was loaded onto a HisTrap FF Crude 5 mL column (GE Healthcare) pre-equilibrated in His buffer A. Protein was eluted in an imidazole gradient with 0-100% His buffer B. Protein eluate was dialysed overnight in NR dialysis buffer (25 mM Tris-HCl pH 8.0) at 4°C in 10 kDa MWCO snakeskin dialysis tubing (Thermo Fisher Scientific) in the presence of 50 µg/mL PreScission™ protease (produced in-house) to cleave off the C-terminal 2StrepII-10His tag. Protein was loaded onto a HiTrap Chelating HP 5 mL column (GE Healthcare) pre-equilibrated in NR dialysis buffer, to separate cleaved protein from the C-terminal tag. Eluted protein was then subjected to anion exchange on a HiTrap Q HP 5 mL column (GE Healthcare) pre-equilibrated in NR Q buffer A (25 mM Tris-HCl and 20 mM NaCl). Protein was eluted in a NaCl gradient with 0-75% NR Q buffer B (25 mM Tris-HCl and 1 M NaCl) and protein eluted at 500-600 mM NaCl. Fractions were analysed by non-reducing 12% SDS-PAGE and relevant peak fractions were pooled and concentrated to 3 mL using Vivaspin-20 10 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein was injected onto a HiLoad 16/60 Superdex 75 prep grade column (GE Healthcare) pre-equilibrated in NR SEC buffer. Peak fractions were analysed by non-reducing 12% SDS-PAGE and then pooled and concentrated to 100-200 µL and 3-6 mg/mL using Vivaspin-6 10 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein concentrations were measured using the Nanodrop 2000c (Thermo Fisher Scientific). Protein was aliquoted in 50 µL aliquots, snap frozen in liquid nitrogen and stored at -80°C.

2.6.4 Refolding recombinant FGF19 and FGF21 proteins from inclusion bodies

Competent C41 (DE3) *E. coli* cells were transformed with a FGF19 (23-216) or FGF21 (29-209) construct in the pAZET09 vector. Glycerol stocks prepared from these were used to inoculate cultures of 500 mL TB medium containing ampicillin (50 mg/L) and cultures were shaken at 37°C. Following induction of expression with 0.4 mM IPTG cells were shaken overnight at 37°C. Cells were then harvested and frozen. Harvested frozen cell pellets were thawed on ice, lysed by chemical lysis and the supernatant discarded. The pellets contained inclusion bodies of FGF19 or FGF21. Pellets were washed sequentially with three solutions buffered with 50 mM Tris-HCl pH 8.0 (centrifuging at 20,000 x *g*, 30 minutes and 4°C between washes and discarding the supernatant): 2% Triton-X100, 1 M NaCl, and just buffer. The purified inclusion bodies were then reduced (25 mM Tris-HCL pH 8.0 and 20 mM TCEP) and denatured (6 M Gdn-HCl pH 8) followed by centrifugation (20,000 x *g*, 30 minutes and 4°C) and the supernatant placed on ice. Cold, solubilised inclusion bodies in the supernatant were swiftly poured into 20-30 volumes of pre-chilled refolding solution (0.5 M L-arginine, 100 mM NaCl, 100 mM Tris-HCl pH 8.0, 2 mM L-cysteine and 0.2 mM L-cystine) and sealed for 1-2 days at 4°C. Refolded protein was dialysed against 25 mM Tris-HCl pH 8.0 and 1 mM NaCl in 3.5 kDa MWCO snakeskin dialysis tubing (Thermo Fisher Scientific), sterile filtered (0.22 µM) to remove refolding precipitates and loaded onto a HisTrap FF Crude 5 mL column (GE Healthcare) pre-equilibrated in His buffer A. Protein was eluted with 100% His buffer B and dialysed against 25 mM Tris-HCl pH 8.0 and 1 mM NaCl in 3.5 kDa MWCO snakeskin dialysis tubing. Protein was then injected onto a HiLoad 26/60 Superdex 200 prep grade column (GE Healthcare) pre-equilibrated in NR SEC buffer. Peak fractions were analysed by reducing 15% SDS-PAGE, pooled and concentrated to 500-1400 µL, 1-7 mg/mL and over 95% purity in most cases (purity estimated by SDS-PAGE analysis) using fresh Vivaspin-20 3 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein concentrations were measured using the Nanodrop 2000c (Thermo Fisher Scientific) before protein was snap frozen in liquid nitrogen and stored at -80°C.

2.7 Production of phosphorylated recombinant FGFR3 KD protein

Recombinant unphosphorylated FGFR3 KD proteins at 2-10 mg/mL were incubated with 25 mM $MgCl_2$ and 10 mM ATP at 22°C for 45 minutes for autophosphorylation to occur. Kinase reactions were stopped with 50 mM EDTA and desalted on a HiLoad 16/60 Superdex 75 column (GE Healthcare) using an Akta Purifier (GE Healthcare) with desalting buffer (25 mM Tris-HCl pH 8.0, 10 mM NaCl and 1 mM TCEP). Peak fractions corresponding to FGFR3 were pooled and loaded on to a 1 mL Resource Q anion-exchange column (GE Healthcare) equilibrated in Q buffer A. Differently phosphorylated

FGFR3 protein was eluted with a very shallow linear gradient over 500 column volumes to 40% of Q buffer B. The resulting chromatogram showed five peaks, corresponding to the 0p, 1p, 2p, 3p and 4p phosphorylated forms of FGFR3. This was confirmed by native PAGE analysis (gels prepared in-house) of pooled fractions from each peak. The 4p peak fractions were pooled and concentrated to 1-2 mg/mL using a Vivaspin-20 10 kDa MWCO centrifugal concentrator (Sartorius Stedim Biotech). Protein was aliquoted in 50 µL aliquots, snap frozen in liquid nitrogen and stored at -80°C.

2.8 Top-down native mass spectrometry

Recombinant purified protein samples were diluted to 0.1 mg/mL in UPLC-grade H_2O containing 0.1% v/v formic acid and 10 µL of each sample was analysed on a Acquity UPLC® system (Waters) using a C4 column coupled to a single-quadrupole mass spectrometer. Samples were subjected to electrospray ionisation in positive ion mode. Absolute molecular masses were determined (to the nearest Dalton) following deconvolution of the mass/charge spectra using a maximum entropy method in-built in the Waters software.

2.9 Kinase assays

In vitro kinase assays were performed at 21-22°C using the ADP-Glo™ kinase assay (Promega). Assay reagents were prepared as per the manufacturer's instructions, and only Ultra Pure ATP (Promega) was used. Kinase reactions were carried out in kinase reaction buffer (KRB) (initially buffered with 40 mM Tris-HCl pH 8.0 during optimisations with FGFR3-2C, and later with HEPES-HCl pH 7.5 following buffer screening, both versions containing 20 mM NaCl, 20 mM $MgCl_2$, 100 µM Na_3VO_4 and 100 µM TCEP) in 15 µL triplicate reactions unless otherwise indicated. All assays in which the final kinase concentrations were predominantly below 0.5 µM also contained 2 mM $MnCl_2$ and 0.1 mg/mL BSA. A poly-Glu-Tyr ($polyE_4Y_1$) peptide (Sigma) was used as a synthetic kinase substrate. All reaction components were serially diluted in KRB. Reactions were started by the addition of either ATP or kinase, depending on the type of experiment. After the required reaction duration, kinase reactions were stopped by the addition of 15 µL ADP-Glo™ Reagent for 40 minutes, followed by the addition of 30 µL Kinase Detection Reagent for 30-60 minutes (depending on the concentration of ATP used in the assay). All 60 µL of each ADP-Glo™ reaction was then transferred to the corresponding well of a white opaque shallow 96-well ProxiPlate (Perkin Elmer), unless otherwise indicated, and luminescence was measured at 520 nm on a FLUOstar Optima microplate reader (BMG Labtech) in arbitrary luminescence units (AU). Data were

analysed using Prism software (GraphPad). Different equations, built-in to the Prism software, were used for non-linear regression (curve fitting) for analysing data from different types of assays. For autophosphorylation and substrate phosphorylation dose response experiments, the relevant saturation dose response binding equations (equations **1** and **2** respectively) were used to fit the data. For time course experiments, a one phase exponential model for pseudo-first order association kinetics (equation **3**) was used to fit the data, and the equilibrium dissociation constant (K_D) concentrations for each FGFR3 variant extracted from these fits were then used to set up the ATP titration experiments. Data from ATP titration experiments were fit using the Michaelis-Menten equation (equation **4**), from which the affinity constant for ATP ($K_{M(ATP)}$) and the maximum velocity (V_{max}) were extracted and used to calculate the turnover rates and catalytic efficiencies for each FGFR3 variant. Error propagation during such calculations was performed using the relevant standard product or quotient rules.

Equation **1**: 'Log(agonist) vs. response' (with variable Hill slope) equation. This was used to fit data from autophosphorylation dose response experiments to improve the visual clarity of displayed data. It assumes that ligands are in sufficient excess such that negligible ligand depletion occurs over the duration of the experiment. Note that although the equation analyses the data using logarithmic-transformed values of FGFR3 concentrations, the data in this work are subsequently displayed in non-logarithmic form.

$$y = Y_{min} + \frac{Y_{max} - Y_{min}}{1 + 10^{H(\log \mathrm{EC}_{50} - x)}}$$

where y – phosphorylation (arbitrary units, AU), x – FGFR3 concentration (μM), Y_{max} – maximum phosphorylation (AU), Y_{min} – minimum phosphorylation (AU), H – Hill slope, EC_{50} – FGFR3 concentration that produces a response half way between Y_{max} and Y_{min} (μM).

Equation **2**. 'One site specific binding' (with variable Hill slope) equation. This was used to fit data from polyE_4Y_1 substrate phosphorylation dose response experiments to improve the visual clarity of displayed data. It assumes that enzymes and ligands are equally accessible to each other, enzymes are free or ligand bound (no intermediate affinity states), ligand binding is reversible and does not alter enzyme or ligands, and all ligands are in sufficient excess such that negligible ligand depletion occurs.

$$y = \frac{Y_{max}\,[x]}{\mathrm{K}_d + [x]}$$

where y – phosphorylation (arbitrary units, AU), $[x]$ – FGFR3 concentration (μM), Y_{max} – maximum phosphorylation (AU), K_d – equilibrium binding constant (μM).

Equation **3**. 'One phase exponential association' equation. This was used to fit data from time course experiments and to calculate the 'half-time' ($T_{1/2}$), the time at which half of the ligand binding sites become occupied with ligand. It assumes that all ligands are in sufficient excess such that negligible ligand depletion occurs, enzymatic ligand binding sites become occupied with ligand over time, fewer ligand binding sites are unoccupied over time so fewer ligands bind and the curve levels off.

$$y = Y_0 + (Y_{max} - Y_0)(1 - e^{-Kx})$$

where y – phosphorylation (arbitrary units, AU), x – time (minutes, min), $Y_0 - y$ when x = 0 (min), $Y_{max} - y$ when x = infinity (min), K – rate constant (inverse minutes, min^{-1}). Also computed $T_{1/2} = \frac{ln2}{K}$ (min).

Equation **4**: 'Michaelis Menten' equation. This was used to fit data from (ATP or $polyE_4Y_1$) substrate titration experiments. It assumes that the rate of production of product is linear over the chosen time period, all ligands (except the one being titrated) are in sufficient excess such that negligible depletion of the secondary ligand occurs, there is negligible product formation in the absence of enzyme, and there are no cooperativity or allosteric effects by either ligands or products that might modulate enzyme activity.

$$v = \frac{V_{max}\,[\mathrm{S}]}{\mathrm{K_M} + [\mathrm{S}]}$$

where v – phosphorylation (arbitrary units, AU), $[\mathrm{S}]$ – ATP or $polyE_4Y_1$ peptide ligand concentration (μM), V_{max} – maximum phosphorylation (AU), $\mathrm{K_M}$ – ligand equilibrium binding constant (μM).

Analyses using equations **1**-**4** also assume that ADP produced in kinase reactions is only produced by the FGFR3 KD variant being screened (no spontaneous ADP generation), that there is no spontaneous degradation or other loss of enzyme over the time course of the reaction, and that there are no effects of cooperativity, allostery or inhibition by substrates or reaction products.

2.10 General mammalian cell culture materials and methods

2.10.1 Maintenance of adherent mammalian cell lines

All cells were maintained in T25 and T75 culture flasks (Corning) at 37°C with 5% CO_2 in a humidified incubator unless otherwise specified. NIH 3T3 cells were cultured with 10% CO_2 in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) (Sigma) supplemented with 10% v/v Foetal Bovine Serum (FBS) (Thermo Fisher Scientific) and

2 mM GlutaMAX™ (Thermo Fisher Scientific). Stable pFB-FGFR3b NIH 3T3 cell lines were cultured similarly with 10% CO_2 in this medium supplemented with a maintenance dose of 100 µg/mL hygromycin B (Invitrogen). Phoenix™ cell lines were cultured in the same medium supplemented with a maintenance dose of 300 µg/mL hygromycin B. HEK 293T cells were cultured in the same medium without any antibiotics, and were transiently transfected when they were 60% confluent. 3T3-L1 pre-adipocytes were cultured in the same medium with 1 x penicillin-streptomycin solution (Sigma). For each cell line, serum free medium (SFM) refers to medium as specified above, but without FBS.

2.10.2 Maintenance of suspension mammalian cell lines

All cells were maintained in vented Erlenmeyer flasks (Corning) in a humidified incubator at 37°C and 8% CO_2 with shaking at 130 rpm (FreeStyle™ 293F) or 5% CO_2 with shaking at 140 rpm (Expi293F™ and CHO-EBNA GS cells, maintained at AstraZeneca by David Fisher). FreeStyle™ 293F cells were cultured in FreeStyle™ 293 expression medium (Thermo Fisher Scientific) and maintained between 0.1-2.0 x 10^6 cells/mL. Expi293F™ cells were cultured in Expi293F™ expression medium (Thermo Fisher Scientific) and maintained between 0.2-2.0 x 10^6 cells/mL. Proprietary CHO-EBNA GS cells were cultured in CD CHO medium (Thermo Fisher Scientific) supplemented with 25 µM methionine sulphoximine and 100 µg/mL hygromycin B (supplements were removed from culture medium from two passages prior to transfection to transfection onwards) and maintained between 0.2-2.5 x 10^6 cells/mL. Expi293F™ and CHO-EBNA GS cells were maintained at AstraZeneca by David Fisher.

2.10.3 Live Cell Imaging and Counting

Images of live cells in culture were captured using the Axiovert 25 microscope (Zeiss) configured for brightfield phase contrast or green epifluorescence microscopy. Where required, cells were counted using the Cellometer® Auto T4 (Nexcelom Bioscience).

2.10.4 Transient Transfections

Adherent HEK 293T cells were transfected with FuGENE HD (Promega) using (per mL medium) 50 µL Opti-MEM™ SFM (Thermo Fisher Scientific), 1 µg purified DNA and 3 µL FuGENE® HD incubated together for 15 minutes at room temperature before adding these complexes to cells. FreeStyle™ 293F cells were transfected with a branched 10-25 kDa polyethylenimine (PEI) cationic polymer (Sigma) using (per mL of cell suspension) 40 µL Opti-PRO™ SFM (Thermo Fisher Scientific) supplemented with 4 mM L-glutamine (Thermo Fisher Scientific), 1.25 µg purified DNA and 1.875 µg PEI incubated together for 10 minutes at room temperature before adding these complexes to cells.

Expi293F™ cells were transfected with a linear 40 kDa PEI Max cationic polymer (Polysciences) using (per mL of cell suspension) 100 µL Expi293F™ expression medium (Thermo Fisher Scientific), 1.5 µg purified DNA and 6 µg PEI Max incubated together for 15 minutes at room temperature before adding these complexes to cells. CHO-EBNA GS cells were transfected by the addition of (per mL of cell suspension) 0.5 µg purified DNA and 7 µg of the linear 40 kDa PEI Max (Polysciences) directly into the cell suspension. Transfection of Expi293F™and CHO-EBNA GS cells was performed at AstraZeneca under the supervision of David Fisher.

2.11 Production of stable NIH 3T3 cell lines by retroviral transduction

Purified plasmid preparations (2.5 µg) of the pFB empty vector and the pFB vector harbouring human FGFR3b wildtype (WT) and mutants were mixed with TransIT®-293 transfection reagent (Mirus) as per the manufacturer's instructions, in the presence of SFM for 20 minutes at room temperature. These complexes were transfected into 50-70% confluent Phoenix™ cells and maintained daily in fresh medium supplemented with 300 µg/mL hygromycin B. After 72 hours the retroviral supernatant was harvested, sterile filtered (0.45 µm) and mixed with 8 µg/mL polybrene (Sigma). NIH 3T3 cells were incubated in the retroviral supernatant for 4 hours, and then maintained in fresh medium until they were considered 'stable' (two passages post-transduction, 72 hours). Successfully transduced cells were propagated by antibiotic selection with fresh medium supplemented with 200 µg/mL hygromycin B before freezing aliquots of the stable cell lines. Upon re-seeding from stocks, cell lines were subjected to selection with fresh medium supplemented with 200 µg/mL hygromycin B for a further three passages before they were fully selected. Cells were then maintained in fresh medium containing the maintenance dose of 100 µg/mL hygromycin B.

2.12 Differentiation of 3T3-L1 pre-adipocytes to adipocytes

3T3-L1 mouse pre-adipocytes were cultured to 80% confluence in a T75 and then distributed evenly between 24 wells of 6-well plates (Nunc) with 3 mL per well culture medium. Cells were confluent after 48 hours, so differentiation was initiated after a further 48 hours (96 hours after seeding cells). Differentiation was initiated (Day 0) by treating the cells with MDI induction medium (medium supplemented with 0.5 mM 3-isobutyl-1-methylxanthine, 5 µg/mL insulin, 1 µM dexamethasone and 1 µM rosiglitazone). Three days later, MDI induction medium was replaced with insulin medium (medium supplemented with 0.2 µg/mL insulin). Three days later, insulin medium was replaced with fresh culture medium. Fresh culture medium was replaced twice more every two

days, at which point cells were considered sufficiently differentiated to be used for experiments (Day 10-13). This method of adipocyte differentiation was modified from a published 3T3-L1 differentiation protocol (Zebisch et al., 2012).

2.13 Anchorage independent growth assays

Stable NIH 3T3 WT and mutant cell lines were seeded in triplicate at 5 x 10^3 cells per well of a 6-well plate (TPP) in medium containing 0.4% w/v agarose (diluted from 4% w/v agarose, Thermo Fisher Scientific), on a base of medium containing 0.8% w/v agarose. Cells were fed weekly with medium containing 0.4% w/v agarose. After two weeks, cells were stained for 24 hours with 0.3% w/v p-iodonitrotetrazolium violet (Sigma) and a representative area of each well was imaged using a Hamamatsu ORCA-285 camera coupled to a Nikon SMZ1000 microscope. Binary masks were applied to the images and thresholding parameters for area and perimeter of colonies with a mean diameter greater than 100 µm were set on ImageJ. Colonies were counted electronically using ImageJ only if they satisfied both the area and perimeter criteria above the threshold values, and colony counts were then manually checked and adjusted if necessary.

2.14 Protein extraction from adherent mammalian cells

Cultured cells were either starved in SFM or left 'unstarved' (as appropriate), and then lysed *in situ* with 75 µL mammalian-lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% v/v Triton X-100, 20 mM sodium pyrophosphate, 25 mM sodium fluoride, 1 mM β-glycerophosphate, 3 mM Na_3VO_4 and 1 µg/ml leupeptin) containing cOmplete™ protease inhibitor cocktail (Roche). Lysates were clarified by centrifugation (16,000 x *g*, 5 minutes and 4°C). Total protein concentration was determined using the Pierce™ BCA Protein Assay kit (Thermo Fisher Scientific).

NIH 3T3 cells were cultured to 70-80% confluence in 2-3 T75 culture flasks per stable cell line. Cultured cells were either starved for 2 hours at 37°C and 10% CO_2 in SFM or were left 'unstarved' in normal medium containing FBS, and then lysed *in situ* with 75 µL (per T75 culture flask) mammalian-lysis buffer. Protein was then extracted and lysates clarified. 3T3-L1 cells were differentiated in 6-well plates as described above. Differentiated cells were then starved (Day 10-13 differentiated adipocytes) overnight in SFM containing 0.2% BSA (Sigma) while partly differentiated cells were left 'unstarved' (Day 0-10 pre-adipocytes) in normal medium containing FBS. Pre-adipocytes were then lysed *in situ* with 75 µL (per well of 6-well plate) mammalian-lysis buffer and protein extracted. Differentiated adipocytes were treated with FGFs at varying concentrations

(or no FGF controls) with 0.2% BSA for 10 minutes before being lysed *in situ* with 75 µL (per well of 6-well plate) mammalian-lysis buffer. Protein was then extracted and lysates clarified.

2.15 Immunoprecipitation of FGFR3 protein from NIH 3T3 cells

Anti-FGFR3 primary antibody F3922 (Sigma) diluted at 1:25 in mammalian-lysis buffer (1 µg antibody per 1 µg of NIH 3T3 lysate) was incubated for 10 minutes at room temperature with magnetic Protein A Dynabeads® (Thermo Fisher Scientific) pre-equilibrated in mammalian-lysis buffer (5 µL Dynabead® suspension per 1 µg F3922 anti-FGFR3 antibody). Antibody:Dynabead® complexes were washed and incubated with clarified NIH 3T3 lysates (prepared as in section 2.14) overnight at 4°C. Target protein-bound antibody:Dynabead® complexes were washed and protein was eluted by boiling in 4 x SDS-PAGE sample buffer. Eluted supernatant was then analysed by immunoblotting. Immunoprecipitation and immunoblots shown in Figure 3.19b were performed by the Research and Clinical Pathology Support Unit at the UCL Cancer Institute.

2.16 Antibody array

A PathScan® intracellular signalling antibody array kit with chemiluminescent readout (Cell Signaling Technologies) was purchased and used as per the manufacturer's instructions, using the array blocking buffer and detection antibody cocktail supplied in this kit and using NIH 3T3 lysates prepared as in section 2.14. Arrays were imaged using the Odyssey® Fc system (Li-Cor Biosciences). Densitometric analyses were performed using Image Studio™ Lite (Li-Cor Biosciences).

2.17 Small-scale expression tests with suspension mammalian cell lines

Cells (10 mL) were seeded in vented 50 mL mini bioreactors (Corning) at densities of 1.0-1.4 x 10^6 cells/mL (FreeStyle™ 293F), 3.5-4.2 x 10^6 cells/mL (Expi293F™) and 1.0-1.2 x 10^6 cells/mL (CHO-EBNA GS). Cells were transiently transfected using the PEI (FreeStyle™ 293F) or PEI Max (Expi293F™ and CHO-EBNA GS) methods. After 24 hours post-transfection, Expi293F™ and CHO-EBNA GS cells were fed with 10 mL Expi293F™ expression medium (Expi293F™ cells) or 6.7 mL CD CHO medium containing a proprietary MedImmune feed 'M20A' (CHO-EBNA GS cells) and were transferred to a humidified incubator at 37°C and 5% CO_2 with shaking at 700 rpm.

For FreeStyle™ 293F expression tests, samples of cells and media were taken every day for 3-4 days post-transfection for all constructs. For Expi293F™ and CHO-EBNA GS transfections samples of cells and media were taken at 3 days (proteins expected to be intracellularly expressed) and 7 days (proteins expected to be secreted into the medium) post-transfection. Cell samples were lysed in 0.1 volumes PBS-lysis buffer (PBS pH 7.2, 1% Triton X-100 and 1 x cOmplete™ protease inhibitor cocktail [Roche]). Media and whole-cell lysate samples were then denatured with 4 x SDS-PAGE sample buffer and analysed for recombinant protein expression by SDS-PAGE and immunoblotting.

2.18 Purification of recombinant FGFR1c proteins from conditioned media from mammalian cell lines

Medium harvested from small-scale expression tests of FGFR1c (22-374) constructs in FreeStyle™ 293F cells (Day 3, 30 mL), Expi293F™ cells (18 mL, Day 7) and CHO-EBNA GS cells (13 mL, Day 7) was loaded onto a HisTrap FF Crude 5 mL column (GE Healthcare) pre-equilibrated in His buffer A. Proteins were eluted with 100% His buffer B and these eluates were injected onto a HiLoad 26/60 Superdex 200 prep grade column (GE Healthcare) pre-equilibrated in NR SEC buffer. Peak fractions were analysed by reducing 10% SDS-PAGE, pooled and concentrated to 400-500 µL, 1-5.5 mg/mL and over 95% purity in most cases (purity estimated by SDS-PAGE analysis) using fresh Vivaspin-20 10 kDa MWCO centrifugal concentrators (Sartorius Stedim Biotech). Protein concentrations were measured using the Nanodrop 2000c (Thermo Fisher Scientific) before protein was snap frozen in liquid nitrogen and stored at -80°C.

2.19 Homology modelling of human βKlotho KL1 and KL2 domains

Human βKlotho KL1 and KL2 domains share just 28% and 36% sequence identity and similarity respectively. Therefore the two KL domains were modelled independently, against two different templates, using the SWISS-MODEL web-based protein structure modelling workspace (Arnold et al., 2006). Each template was optimally selected from searches against the SWISS-MODEL template library (up-to-date as of 24th June 2017). Human βKlotho KL1 (53-508) and KL2 (517-967) sequences were used as input queries for template searches. SWISS-MODEL was then used to build homology models per domain using the top three templates with the 'best match' to each KL domain (based on sequence identity, SWISS-MODEL GMQE and QMEAN scores for model quality evaluation and filtering for resolution of the PDB structure <2.2 Å). Finally, the model with the highest GMQE and QMEAN model quality evaluation scores for each KL domain was selected for further analysis.

2.20 Biophysical techniques for analyses of recombinant proteins

2.20.1 Thermofluor assay

Recombinant FGFR3 KD protein samples were mixed with Sypro® Orange fluorescent dye (Thermo Fisher Scientific) at final concentrations of 1 mg/mL protein and 1:500 dye respectively in triplicate reactions. Samples were then subjected to a temperature gradient between 10-95°C increasing at 0.5°C intervals (10 seconds each) on a MyiQ™ Real-Time PCR detection system (Bio-Rad). Data were analysed using Prism software (GraphPad) with a custom Boltzmann sigmoid equation with linear baselines.

2.20.2 Circular Dichroism (CD) spectroscopy

Purified FGF proteins were diluted to 200 µL in NR SEC buffer to 1 mg/mL and then dialysed against 3 x 2 L CD buffer (20 mM sodium phosphate pH 7.5) changed over 24 hours at 4°C. Final protein concentrations were re-measured using the Nanodrop 2000c (Thermo Fisher Scientific) before CD experiments were performed on the Jasco J-720 instrument using a 0.1 mm path length cuvette. Ten scans were taken for each protein (using CD buffer as a reference) between 180-300 nm at 1 nm intervals. Scans were averaged, reference subtracted and zeroed using CDtool software (Lees et al., 2004). Secondary structure analyses of the processed data were performed on the online DichroWeb server using the CDSSTR algorithm (Manavalan and Johnson, 1987) with reference dataset SP175 (Lees et al., 2006) as well as the online tool BeStSel (Micsonai et al., 2015) on the highest quality data between 190-240 nm.

2.20.3 Isothermal Titration Calorimetry (ITC)

Recombinant purified FGF and FGFR1c proteins were buffer exchanged into the same batch of NR SEC buffer using a HiLoad 26/60 Superdex 200 prep grade column (GE Healthcare). All proteins were further diluted using this same batch of NR SEC buffer where necessary. A MicroCal™ iTC-200 instrument (MicroCal) was used. For all experiments, FGF ligands (200 µM) were injected from the syringe into the sample cell containing FGFR1c protein (20 µM) purified from FreeStyle™ 293F cells. Experiments were all performed at 25°C. Data were analysed using Origin® software (MicroCal) by fixing the reaction stoichiometry as 1 (the theoretical stoichiometry for this interaction) and attempting to fit the data to a one-site model within the software.

2.20.4 Surface Plasmon Resonance (SPR) spectroscopy

Experiments were conducted on the Biacore™ T200 instrument (GE Healthcare) using NTA sensor chips with four flow cells (Fc) per sensor chip. FGF1, FGF19 and FGF21 (and ‘no FGF’) were diluted from frozen stocks to 1-2 µg/mL in HBS-EP buffer (10 mM

HEPES-NaOH pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.05% v/v Tween® 20). Diluted FGFs were immobilised on an NTA sensor chip, using Ni^{2+}-NTA affinity immobilisation followed by primary amine coupling and ethanolamine termination, and using a 30 µL/minute flow rate with a target immobilisation level of 30 response units (RU). A two-fold dilution series of purified FGFR1c protein was prepared in HBS-EP buffer and kinetics experiments were performed with these analytes using a 30 µL/minute flow rate. Analyte was injected for 180 seconds (association) and then HBS-EP buffer was injected for 180 seconds (dissociation). The chip surface was washed with 10 mM glycine pH 3.0 following injection and regenerated with 2 M NaCl. Affinity calculations were performed with the Biacore T200 evaluation software (GE Healthcare).

Chapter 3: Assessment of FGFR cancer-associated mutations

3.1 Introduction

A decade of efforts dedicated to the deep sequencing of patient tumour samples and cancer cell lines has resulted in a wealth of mutation data being accumulated for cancer-associated genes such as FGFRs. As mentioned in Chapter 1, FGFRs are known oncogenes and several types of aberrations affect the FGFR genes giving rise to deregulated signalling in all types of malignancies. Aberrations in FGFRs are estimated to be present in about 7% of solid tumours (Guagnano et al., 2012, Helsten et al., 2016). In particular, cancer-associated somatic point mutations make up almost 30% of FGFR aberrations and have been reported in all four FGFRs and in all types of cancers (Figures 3.1 and 3.2). Components of FGFR signalling are in fact thought to be the most enriched for non-synonymous point mutations in cancer (Greenman et al., 2007). The highest occurrence of point mutations in FGFRs is in FGFR3 followed by FGFR2 and subsequently these are found most commonly in urothelial (bladder) and endometrial carcinomas respectively (Guagnano et al., 2012, Abbosh et al., 2015, Carter et al., 2015, Helsten et al., 2016). Despite the observation that FGFRs (like other RTKs) are frequently found mutated in cancers, development of treatments targeting FGFRs in the clinic have lagged behind those targeting other RTKs, and currently there are no approved agents for the treatment of cancers based on FGFR targeting. Further advancement of the molecular understanding of FGFR function and activation, and the impact of cancer-associated mutations on these phenomena, will aid the development of targeted anti-FGFR agents and will directly inform ongoing clinical trials.

This Chapter will explore the diversity of non-synonymous point mutations in FGFRs in cancer. The functional impact and possible activation mechanisms of commonly occurring, known activating missense substitutions will be discussed with a focus on mutations in FGFR3. Finally, the discovery of acquired mechanisms of resistance in FGFRs (including through point mutations) in cancer in response to clinical treatments and their mechanisms will be presented.

3.1.1 Identification of somatic missense substitutions in FGFRs in cancer

The first identification of a somatic mutation in FGFRs in cancer was in FGFR3 in multiple myeloma (Chesi et al., 1997). In the two decades that followed, thousands of somatic mutations have been identified in FGFR3 and in the other FGFRs through a combination of large-scale cancer genome sequencing projects and smaller lab-based

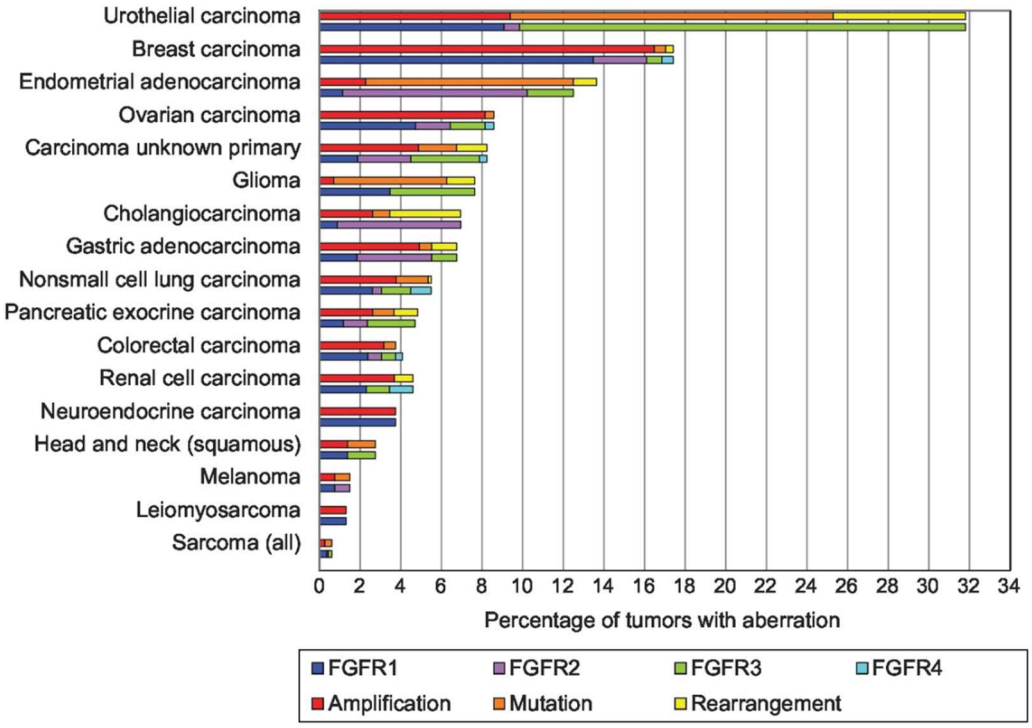


Figure 3.1 Landscape of aberrations in FGFRs in all types of cancers. Bar chart showing the approximate percentages of FGFR alterations found in diverse tumour types in all four FGFRs, including *FGFR* gene amplifications, point mutations and chromosomal rearrangements leading to gene fusions, ranked by the cancer type. The upper bar for each tumour type shows the type of mutation, while the lower bar shows the total percentage of aberrations occurring in each FGFR isoform. Figure from (Helsten et al., 2016).

studies. Two of the largest international consortia (still ongoing) are The Cancer Genome Atlas (TCGA) [cancergenome.nih.gov] (Cancer Genome Atlas Research et al., 2013) and the International Cancer Genome Consortium (ICGC) [icgc.org] (International Cancer Genome et al., 2010) which have currently sequenced more than 11000 and 17000 cancer genomes respectively from 33 cancer types (October 2017). Other smaller studies such as one mentioned in Chapter 1 which analysed nearly 5000 tumour samples specifically to assess the mutation status of FGFRs (Helsten et al., 2016) combined with data from larger collaborative international projects have fostered efforts to comprehensively curate mutation data detected across all studies into central online databases. The largest such database, whose entire integrated dataset is also publicly available, is the Catalogue of Somatic Mutations in Cancer (COSMIC) [cancer.sanger.ac.uk] (Forbes et al., 2015). This database is updated every few months, curating new genes and ensuring the dataset is maintained and kept up-to-date by including new data from the most recent studies. It has provided invaluable insights into patterns of cancer-associated aberrations across different genes and cancer types, including somatic point mutations in FGFRs.

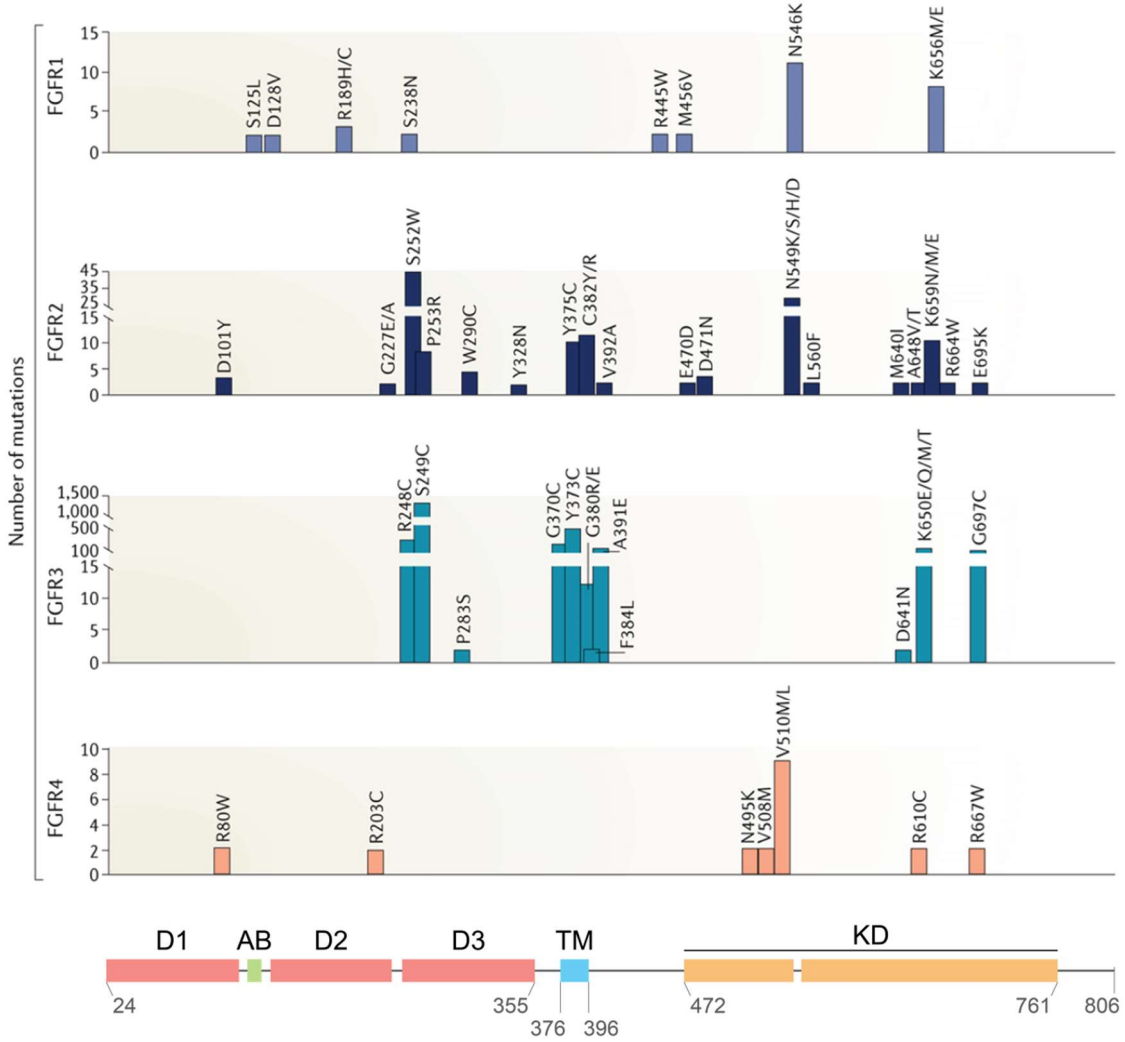


Figure 3.2 Somatic cancer-associated point mutations span the entire length of FGFRs. Total number of observations of somatic point mutations reported in tumours from cancer patients in FGFR1-4 from COSMIC, including insertions and deletions (in-frame and frame shift) and missense substitutions (upper four panels). The most commonly observed mutations (hotspots) in each FGFR are labelled with residue numbering according to the canonical FGFR c splice isoforms. The FGFR domain organisation is illustrated below the mutation histograms using FGFR3c residue numbering as a reference (lower panel), showing the D1, D2 and D3 extracellular immunoglobulin-like domains, acid box (AB), transmembrane helix (TM) and the 'split' tyrosine kinase domain (KD). This illustration corresponds to the locations of mutations in the upper four panels found in each FGFR domain. Figure adapted from (Babina and Turner, 2017).

Rather than being limited only to regions that are important for kinase function (as is the case in closely related RTKs) such as the kinase domain (KD), somatic point mutations in FGFRs have been found to span the entire length of the receptors (Figure 3.2). Notably, several of these mutations are also the same as mutations found in FGFRs in germline disorders such as skeletal dysplasias (Wilkie, 2005, Gallo et al., 2015) and these mutations were discovered in FGFRs in the context of skeletal disorders in the early 1990s, first in FGFR3 (Rousseau et al., 1994, Shiang et al., 1994) and very soon after in FGFR1 and FGFR2 (Jabs et al., 1994, Muenke et al., 1994, Reardon et al., 1994).

The functional impact and mechanisms of action of several of the most frequently occurring point mutations in FGFRs have therefore been investigated over the years in the context of both germline disorders and cancer. What is known so far about these is presented in the next sections with a focus on point mutations in FGFR3.

3.1.2 Cancer-associated mutation hotspots in FGFR3

By far the most frequently detected somatic missense substitution in all FGFRs in cancer is FGFR3-S249C (Figure 3.2) with over 2000 observations to date. This mutation occurs in the D2-D3 linker of the FGFR3 ectodomain, adjacent to another of the 'top 5' most common FGFR mutations, FGFR3-R248C (over 400 observations). Another very frequently detected mutation is FGFR3-Y373C (over 600 observations) which lies in the extracellular juxtamembrane region. These three FGFR3 mutations constitute more than 85% of all cancer-associated mutations detected in FGFR3 to date (di Martino et al., 2016). The other most frequent mutation 'hotspots' in FGFRs (each with 40-500 observations in cancer) all occur in FGFR3, either in the extracellular juxtamembrane or transmembrane region (G370C and A391E) or in the KD (K650E/M and G697C). Similar mutations in these regions also occur in the other FGFRs (Figure 3.2), some at the precise corresponding residues. It should be noted that some mutations were detected in the FGFR b splice isoforms (depending on the tissue of origin which determines which splice isoform is expressed, see section 1.4), but in this work all mutations will be referred to using the residue numbering of the canonical FGFR c splice isoforms.

3.1.3 Functional impacts of gain-of-function mutations in FGFR3

FGFR3 mutations are found in 75% of non-muscle invasive (lower grade) urothelial carcinomas and in approximately 15% of higher grade muscle invasive urothelial carcinomas (Cappellen et al., 1999, Billerey et al., 2001, Kimura et al., 2001, van Rhijn et al., 2001, Helsten et al., 2016) as well as in about 5% and 3% of cervical and small cell lung carcinomas respectively (Rosty et al., 2005, Helsten et al., 2016). Therefore, many of the point mutations listed above have been rigorously studied to assess their functional impact in the context of cancers and germline disorders by assessing their effects on FGFR dimerisation and activation. Such experiments have helped classify mutations into driver mutations that might contribute directly to development and/or maintenance of disease, or passenger mutations that might just be passive and not play direct roles in pathogenesis. It is thought that many of these 'drivers' are gain-of-function mutations that constitutively activate FGFRs, allowing illicit ligand-independent signalling, and evidence for this is discussed in the next sections.

3.1.4 Oncogenic FGFR3 extracellular and transmembrane region mutations

Several frequently observed FGFR3 variants in the ectodomain or the transmembrane region result in the introduction of a new cysteine residue, particularly R248C, S249C, G370C and Y373C. This led to the hypothesis that the mutations may activate FGFRs by constitutive dimerisation of the extracellular or transmembrane domains (Figures 3.3 and 1.12c). Early studies supported this hypothesis by showing (in NIH 3T3, COS or BaF3 model cell lines) that these mutations were dominantly acting and activating compared to FGFR3 wildtype (WT); they induced ligand independent dimerisation through disulphide bonding (abrogated upon chemical reduction), activation and constitutive tyrosine autophosphorylation *in vitro* (Naski et al., 1996, d'Avis et al., 1998, Adar et al., 2002, Hafner et al., 2010). Subsequently, studies focusing on S249C and Y373C showed that these mutations morphologically transform NIH 3T3 cells, hyperactivate several downstream signalling pathways, promote cell proliferation, increase cell saturation density and permit anchorage independent growth, and additionally these effects could be reversed by targeted gene knockdown or treatment with FGFR inhibitors (including in relevant bladder cancer cell lines expressing the FGFR3 mutants) (Chesi et al., 2001, Ronchetti et al., 2001, Bernard-Pierrot et al., 2006, Tomlinson et al., 2007b, di Martino et al., 2009, Qing et al., 2009, Chell et al., 2013, Williams et al., 2013). Furthermore, S249C (Bernard-Pierrot et al., 2006) and Y373C (Trudel et al., 2004) also produced tumours in xenograft models in nude mice and tumour growth could be inhibited *in vivo* by FGFR3 inhibitors (Trudel et al., 2004, Xin et al., 2006, Miyake et al., 2010), providing further evidence for the drug sensitivity and oncogenicity of these mutations. Similar studies performed for R248C and S249C in cellular and xenograft lung cancer models showed that these mutations are sensitive to FGFR inhibitors and the tumours are oncogene-addicted to mutant FGFR3 (Liao et al., 2013). The oncogenic effects and sensitivity of these activating FGFR3 mutants to FGFR inhibitors strongly implicates them as oncogenic drivers.

However, despite the notion (supported by evidence from these cellular studies) that these cysteine mutants constitutively phosphorylate and activate FGFRs through ligand independent dimerisation, many of these studies suggested that this might only be a partial effect, with some less activating mutants (e.g. Y373C) becoming fully activated in a partially ligand dependent manner. Additionally, Förster Resonance Energy Transfer (FRET) experiments suggest that the R248C, S249C and Y373C mutations only partially stabilise FGFR3 dimers, but might instead activate FGFRs by inducing conformational perturbations (rather than constitutive dimerisation) that juxtapose the dimers in the correct orientation for full activation (Del Piccolo et al., 2015) while less activating mutations might only achieve correct juxtaposition (for full activation) upon ligand binding

(Adar et al., 2002). This mechanism of activation has also been proposed for the transmembrane non-cysteine activating hotspot mutation A391E suggesting it traps the dimer in the active conformation, mimicking the ligand-bound state (Chen et al., 2013a, Sarabipour and Hristova, 2015, Sarabipour and Hristova, 2016). Furthermore, many of the studies presented here have suggested that the position of the mutated residue in the extracellular or transmembrane region might affect the level of FGFR activation in a 'graded' manner. This might lead to differential activation of the receptors and modulated signalling that presents as lower or higher grade tumours. The precise molecular activation mechanisms of these mutations remain to be fully elucidated, as is necessary for the development of precision targeted therapies against subsets of tumours harbouring these different mutations.

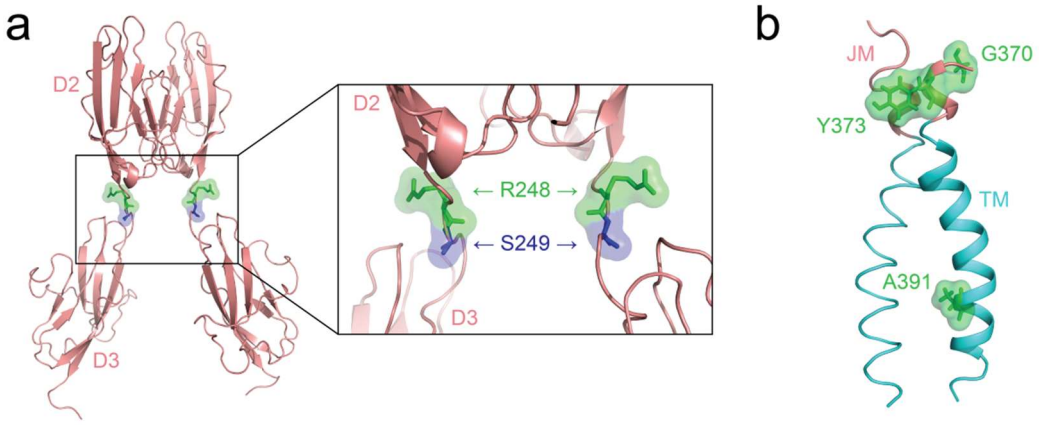


Figure 3.3 Locations of hotspot mutations in the extracellular and transmembrane domains of FGFR3. The cysteine mutations are thought to constitutively dimerise, and thus activate, FGFR3. **(a)** Structural location of the two most frequently observed mutation hotpots in the FGFR3 extracellular region (R248C and S249C, shown as sticks against a surface and coloured green and navy blue respectively) mapped on a cartoon representation of the crystal structure of the FGFR3c extracellular region dimer [PDB: 1RY7] (Olsen et al., 2004) (coloured salmon), showing that they occur in the D2-D3 linker (left panel). A close-up view of this linker is shown in the right panel. **(b)** Structural location of the three most frequently observed mutation hotspots in the FGFR3 transmembrane (TM) region (A391E) and extracellular juxtamembrane (JM) region (G370C and Y373C) mapped on a cartoon representation of the structure of the FGFR3c TM domain dimer [PDB: 2LZL] (Bocharov et al., 2013) (coloured cyan). All three TM and JM hotspot residues are shown as sticks against a surface and coloured green.

3.1.5 Mutation hotspots in the FGFR3 tyrosine KD

Two hotspots in the FGFR3 KD at which mutations are observed most frequently are highlighted in Figure 3.4 (middle panel), namely K650 and G697. Mutations to cysteine at position G697 (which is in the αF-αG loop in the C-lobe of the KD) were all identified in a single study in oral squamous cell carcinomas (OSCC) (Zhang et al., 2005). This mutation is unlikely to activate FGFRs through creating disulphide linked dimers because

of the reducing nature of the cytoplasmic compartment in which the mutation exists. Although this study suggested that G697C is constitutively activating, a follow up study could not detect this mutation in any cases of OSCC (Aubertin et al., 2007), and so the relevance (and/or activation mechanism) of this hotspot mutation in cancer remains unresolved.

In contrast, the activating nature of mutations at position K650 in the A-loop of the KD has been comprehensively demonstrated and the probable mechanisms of activation deciphered. K650E was initially identified in the context of thanatophoric dysplasia type II (Tavormina et al., 1995) where it was shown to be constitutively autophosphorylated, to hyperactivate downstream signalling, and (in most cases) increase proliferation in a (partially) ligand independent manner compared to WT (Naski et al., 1996, Webster et al., 1996, d'Avis et al., 1998, Tavormina et al., 1999, Lievens and Liboi, 2003, Krejci et al., 2008). Subsequently, it was shown to exert the same effects in the context of cancer, where it also transformed NIH 3T3 cells, promoted anchorage independent growth in soft agar, and led to tumour formation in nude mice (Hart et al., 2000, Chesi et al., 2001, Ronchetti et al., 2001, Kong et al., 2002, Agazie et al., 2003). However, unlike cysteine mutations in the extracellular and transmembrane regions, K650E did not induce formation of ligand independent dimers, and appeared to activate FGFR independently of dimerisation (Naski et al., 1996, d'Avis et al., 1998, di Martino et al., 2009). It was also found that the exact replacement at position K650 was important in determining the level of kinase activation, with K650E and K650M being most activating followed by other replacements (Webster et al., 1996, Bellus et al., 2000, Chen et al., 2013b). Additionally, studies in which the A-loop 'twin' tyrosine residues necessary for canonical KD activation were mutated to phenylalanine (Y647F and Y648F in FGFR3) showed that K650E was still activating in this context (Webster et al., 1996, Chen et al., 2005, Huang et al., 2013b). It was therefore proposed that K650E might activate the KD by functionally mimicking the active kinase conformation independently of ligand binding or autophosphorylation.

Recently, the structure of the FGFR3-K650E KD was solved by crystallography, providing molecular evidence for this proposed mechanism of KD activation (Huang et al., 2013b). This structure showed that mutation of position 650 to glutamic acid functionally mimics phosphorylation of the A-loop 'twin' tyrosines to stabilise the active KD conformation through altering two key networks of hydrogen bonds (Figure 3.4). One network is stabilising and is created at the A-loop, between E650 and Y648 and adjacent residues from the C-loop (R616 from the HRD motif) and A-loop (R640, T651 and T652), rendering the A-loop in an identical active conformation (Figure 3.4, right panel) as seen

for phosphorylated, active FGFR1 and FGFR2 KDs (Chen et al., 2007b, Bae et al., 2009). The second network (the 'molecular brake' in the kinase hinge region) stabilises the *inactive* conformation, and E650 allosterically disrupts these autoinhibitory interactions (Figure 3.4, left panel). Recently, it was proposed that the assembly of the A-loop network upon phosphorylation is 'sensed' by the 'DFG-latch' and the 'αC tether' leading them to allosterically and dynamically 'transmit' the signal towards the kinase hinge to disassemble the 'molecular brake' and allow dynamic activation of the kinase (Chen et al., 2017). K650E was proposed to act by 'hijacking' this allosteric mechanism, shifting the equilibrium in favour of the active conformation. In addition, the crystal structure shows the presence of the β10/12 strands and ordering of the C- and R-spines, characteristic of an active KD (Figure 3.4, middle panel). This evidence for dimerisation- and ligand-independent FGFR hyperactivation by K650E strongly implicates it as a potential oncogenic driver.

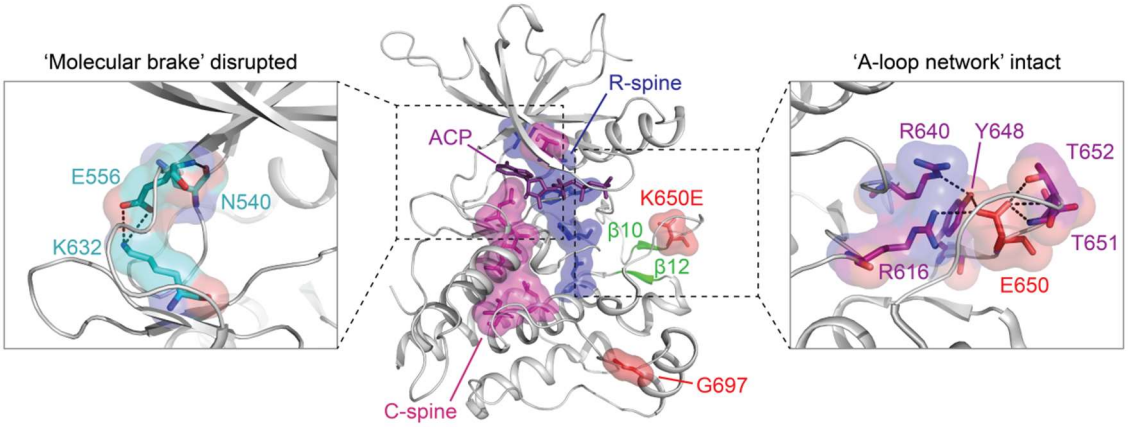


Figure 3.4 Locations of hotspot mutations in the FGFR3 KD and mechanism of activation of pathogenic mutation K650E. Structural locations of the two most frequently observed mutation hotspots in the FGFR3 KD (K650E and G697C, shown as sticks against a surface and coloured red) mapped on a cartoon representation of the crystal structure of the active FGFR3 KD harbouring the K650E mutation (coloured grey) and with a non-hydrolysable ATP analogue (ACP, shown as sticks and coloured purple) bound in the ATP binding pocket [PDB: 4K33] (Huang et al., 2013b), showing that K650E is found in the A-loop and G697C is found in the αF-αG loop (middle panel). The mechanism of KD activation by K650E is depicted on this structure in the three panels shown in the same orientation. The middle panel shows ordering of the hydrophobic regulatory and catalytic (R- and C-) spines (residues shown as sticks against a surface and coloured navy blue and bright pink respectively) and presence of the β10/12 strands (coloured green). The left panel is a close-up view of the kinase hinge region showing disengagement of the hydrogen bonding network formed by the triad of 'molecular brake' residues (shown as sticks against a surface and coloured cyan and by atomic element with nitrogen atoms in navy blue and oxygen in red). The right panel is a close-up view of the A-loop showing an intact hydrogen bonding network formed between residues R616 from the C-loop and R640, Y648, E650, T651, and T652 from the A-loop (shown as sticks against a surface and coloured purple and by atomic element with nitrogen atoms in navy blue and oxygen in red, except for E650 which is coloured red). Hydrogen bonds in left and right panels are shown as dashed lines and coloured black.

3.1.6 Acquired resistance through FGFR mutation and/or RTK co-activation

The phenomenon of acquired resistance mutations in RTKs in response to targeted therapies was introduced in section 1.1.2. Resistance mutations arise in response to treatment with small molecule inhibitors that bind to the ATP binding site, so resistance mutations tend to occur in the ATP binding pocket in the KD. In FGFRs, the crucial gatekeeper residue is normally a valine (unlike threonine in most RTKs) and the first gatekeeper mutation identified in FGFRs as an acquired resistance mutation was FGFR3-V555M in a human cancer cell line resistant to FGFR inhibitors (Chell et al., 2013). This mutation and mutations at the corresponding gatekeeper residues in FGFR1 and FGFR2 (V561M and V564I respectively) have been investigated in pre-clinical models. They exhibit variable (and currently slightly contradicting) resistance and sensitivity to different multi-kinase and FGFR-selective inhibitors in cellular and *in vitro* activity assays (Byron et al., 2013, Nakanishi et al., 2014, Bunney et al., 2015, Sohl et al., 2015). However, all these studies demonstrate that gatekeeper mutations exhibit elevated kinase activity compared to WT, as also shown for other RTKs (Azam et al., 2008) and are predicted to shift the equilibrium to the active KD conformation by stabilising the R-spine (Figure 3.5a). Gatekeeper mutations therefore seem to function by a dual mechanism involving kinase activation as well as steric hindrance of inhibitor binding to the ATP binding pocket (Figure 3.5b), imparting resistance to first-generation ATP-competitive targeted therapies. Other mutations in the kinase hinge, 'molecular brake' and A-loop of FGFR2 (some of which are activating) have also been shown to exhibit different levels of resistance to kinase inhibitors in cellular studies (Byron et al., 2013). These include mutations at the corresponding positions to FGFR3 N540 and I538 (highlighted in Figure 3.5b), although these mutations have not yet been observed as acquired resistance mutations in response to clinical targeted inhibition. Recently, second-generation inhibitors have been developed that can irreversibly bind to the active site, some of which are able to bind and inhibit FGFRs with gatekeeper and other activating mutations and inhibit the proliferation of cancer cell lines harbouring FGFR aberrations (Tan et al., 2014, Huang et al., 2015). Further development of second-generation inhibitors might be used in the future in the clinic to overcome inhibitor resistance imparted by mutations acquired in response to first-generation targeted therapies.

Acquired resistance mutations can also arise in response to therapeutic targeting by an entirely different mechanism. As mentioned previously, different RTKs exhibit crosstalk between members including FGFRs, and so RTKs are commonly found co-activated in cancers (Xu and Huang, 2010). This results in the formation of multimodal signalling networks for cancer cells to maintain a robust signal with diverse outcomes using a

limited combination of intracellular signalling effectors. Upon inhibition of one RTK with a targeted inhibitor, oncogene switching can occur as a compensatory bypass or 'escape' resistance mechanism to reprogram the kinome through activation of other RTKs, rendering monotherapy ineffective against cancers in which RTKs are co-activated. FGFR signalling has been found co-activated with several other RTKs in cancers such as gliomas (Stommel et al., 2007, Pillay et al., 2009), cervical carcinoma (Pietras et al., 2008), lung cancer (Ware et al., 2010, Cascone et al., 2011) squamous carcinoma (Ciaccio et al., 2010), breast cancer (Stuhlmiller et al., 2015), and in FGFR-addicted urothelial, breast and gastric cancer cell lines (Wilson et al., 2012). Interestingly, in one patient that developed resistance in response to treatment with EGFR inhibitors, it was found that a mutation at one of the A-loop 'twin' tyrosines in the FGFR3 KD (Y647C) might be the cause of resistance through oncogene switching to FGFR signalling (Crystal et al., 2014). Although not yet experimentally demonstrated, it was predicted that this mutation would be likely to be activating due to its location close to the FGFR3-K650 mutation hotspot.

Several studies have shown that in cancer cells in which FGFRs are co-activated with other RTKs (either as a response to inhibition of the other RTK, or when another RTK is co-activated in response to FGFR-targeting), monotherapy is often ineffective and only combination therapy with more than one RTK inhibitor can produce the synergistic cancer-killing effect to re-sensitise cancer cells to chemotherapy (Ezzat et al., 2005, Herrera-Abreu et al., 2013, Singleton et al., 2013, Terai et al., 2013, Holdman et al., 2015, Kim et al., 2015, Kim et al., 2016, Wong et al., 2016, Hanker et al., 2017). Cancers that are addicted to FGFR signalling are more likely to be sensitive to FGFR-targeted therapies, making FGFRs important therapeutic targets in such cancers. Therefore understanding FGFR aberrations in cancer, their driver or addiction effects on tumours, and their effects on the efficacy of targeted therapeutics is crucial before combination therapies against FGFRs and other RTKs in cancer can be successfully employed in the clinic.

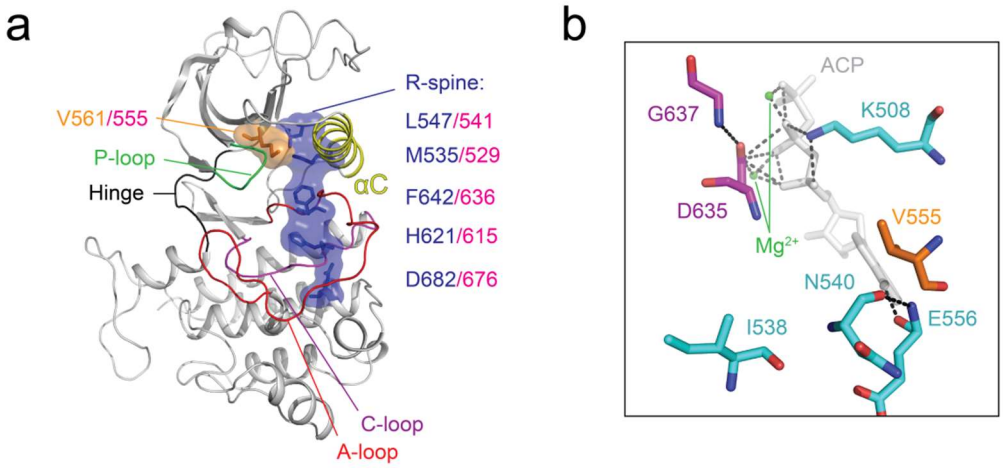


Figure 3.5 Mechanism of activation by gatekeeper mutations in the FGFR KD. (a) Cartoon representation of the crystal structure of the active FGFR1 KD (coloured grey) harbouring the V561M gatekeeper mutation (shown as sticks against a surface and coloured orange) [PDB: 4RWI] (Sohl et al., 2015). The αC-helix (yellow) P-loop (green), kinase hinge (black), C-loop (magenta) and A-loop (red) are highlighted and labelled. The A-loop is in the active conformation and ordering of the regulatory R-spine is shown, with R-spine residues labelled (shown as sticks against a surface and coloured navy blue). The positions of corresponding residues in FGFR3 are also labelled (bright pink) for the gatekeeper and R-spine residues. **(b)** Illustration of key residues contributing to the ATP binding pocket of FGFR3 shown using a cartoon representation of the only available FGFR3 KD crystal structure (FGFR3-K650E) [PDB: 4K33] (Huang et al., 2013b) in a different orientation to **a**. Non-hydrolysable ATP analogue (ATP, shown as sticks and coloured grey) and magnesium atoms (shown as spheres and coloured green) are shown bound in the ATP binding pocket, which ATP-competitive inhibitors also target. Key residues of the 'molecular brake (cyan), DFG motif in the C-loop (magenta) and the gatekeeper residue (orange) are shown as sticks and coloured by atomic element with nitrogen atoms in navy blue and oxygen in red. Hydrogen bonds are shown as dashed lines and coloured black. The position of V555 in the pocket suggests why mutation to a bulkier residue at this position might sterically hinder ACP (and by extension ATP and ATP-competitive inhibitors) from binding in this pocket.

3.2 Aims

A deeper understanding of the functional impact of FGFR cancer-associated mutations on kinase activation and inhibitor efficacy is needed for better linking of the mutation data to treatment options for patients with the aim of a future personalised approach to FGFR targeting in the clinic. Most studies have so far focussed on understanding the functional impact and mechanisms of action of hotspot mutations, particularly those in the FGFR3 and FGFR2 extracellular and transmembrane regions. Few mutations in the FGFR KD have been rigorously assessed for their functional impact and oncogenic status, with the few studies to date focussing on mutations in the 'molecular brake' and A-loop, and a few studies have analysed similar mutations in the context of skeletal dysplasias. Therefore, there is a need for a direct, comparative, and functional study of cancer-associated FGFR point mutations to differentiate oncogenic driver mutations from passenger mutations in the wealth of genetic mutation data. Additionally, since the majority of targeted FGFR inhibitors currently undergoing clinical trials target the ATP binding pocket in the FGFR KD, it is important to understand the mechanisms of action of mutations in this domain when considering inhibitor binding. Therefore, the diverse, mutation-specific mechanisms affecting kinase activation and ligand binding need to be elucidated to inform better patient selection and increase the success of personalised FGFR targeting in the clinic.

In this work a large panel of representative FGFR mutations from across the KD was assembled using the FGFR3 KD as a model. Purified WT and mutant KD proteins were directly compared through detailed and quantitative measurements by developing *in vitro* kinase activity assays to assess the impact of mutations on kinase activation, ATP binding, turnover and catalytic efficiency. These assays identified potential oncogenic driver mutations and assessed possible mechanisms involved in mutation-specific KD activation. The impact of mutations on the thermal stability of the FGFR3 KD was then assessed using an *in vitro* biophysical assay to further probe possible contributions to KD activation mechanisms. Finally, the functional impact of mutations was compared using complementary experiments in a cellular context to attempt to elucidate mutation-specific impacts on oncogenic FGFR signalling and cellular transformation. These data therefore contribute towards an understanding of the molecular mechanisms of FGFR activation by cancer-associated point mutations through direct impacts on FGFR activity and signalling, and challenge the generalised assumptions of the contributions of hotspot and non-hotspot mutations to KD activation and oncogenesis.

3.3 Results

3.3.1 Database searching and selection of missense substitutions in FGFR3

Somatic cancer-associated missense substitutions have been detected in patient samples along the entire length of FGFR3 and other FGFRs, as discussed in section 3.1. Oncogenic driver mutations in the extracellular and transmembrane regions of FGFRs have been rigorously studied and mechanisms of action investigated (see section 3.1). The focus of this work was to compare and analyse cancer-associated FGFR missense substitution mutations that occur in the cytoplasmic region, particularly in the KD. FGFR3 was used as the model FGFR for this work because missense substitutions occur most frequently in FGFR3, including within this region (Figure 3.2).

As a starting point for this work, cytoplasmic FGFR3 point mutations were sought from COSMIC, the most comprehensive and publicly available database of curated cancer-associated mutations. Of the 63 distinct non-synonymous mutations in the cytoplasmic domain of FGFR3 for all cancer types in COSMIC at the start of this work (November 2013), a comprehensive subset of more than 20 mutations was manually selected for this study. Mutations were selected if they were observed with a high frequency in FGFR3 or if they occurred in functionally important regions or at conserved residues, for example in the A-loop or 'molecular brake' in the KD. Some mutations that occur at low frequencies, even those with only one observation, were also selected to assess whether rare mutations might be relevant in cancer. Since the sequence identity and similarity of the four FGFR KDs is over 80% (see Table 1.3), mutations in COSMIC for FGFR1, FGFR2 and FGFR4 were also accessed and linked to the corresponding residues in FGFR3 based on a multiple sequence alignment of FGFR1-4 and some of these mutations were added to the mutation panel for experimental analysis if they occurred across more than one FGFR or were frequently observed in the specific FGFR isoform. Entries in COSMIC are updated when a new version is released approximately every three months, so more mutations were manually selected for analysis (and others removed) throughout the project until November 2014 (v71) when no further amendments were made to the mutation panel. Mutations present in COSMIC v71 in FGFR3 (and corresponding mutations in the other FGFRs) are listed in Table 3.1 according to their location in the KD.

The compiled data are represented in Figure 3.6a-b as histograms with residue numbering according to FGFR3c and the positions of mutations annotated on the structure of the FGFR3 KD. Analysis of this dataset revealed an evident conservation of mutations across FGFR isoforms, with the corresponding residue in other FGFRs often

also found mutated similarly to FGFR3. The analysis also showed that there are two hotspot positions in the FGFR3 KD at which the highest numbers of mutations are observed (K650 and G697), but revealed a third hotspot in FGFRs (corresponding to FGFR3-N540) which only became apparent when mutation frequencies were pooled across FGFR1-4. The most frequently mutated hotspot was the position corresponding to FGFR3-K650 where mutations to glutamic acid and methionine were much more frequently observed than the less common mutations to asparagine, glutamine and threonine. At the 'pooled' FGFR hotspot corresponding to FGFR3-N540, mutations to lysine were most frequent compared to the less common serine, histidine and aspartic acid substitutions. The final hotspot mutation, G697C, was only observed in FGFR3, although the corresponding mutation to serine was observed rarely in FGFR1 (Table 3.1).

Interestingly, as mentioned previously, somatic mutations found in FGFRs in cancer often correspond to germline mutations observed in developmental syndromes such as skeletal dysplasias (Cappellen et al., 1999, Kimura et al., 2001, Wilkie, 2005, Gallo et al., 2015, Helsten et al., 2015). The positions at which germline mutations are found most commonly in FGFR3 and other FGFRs in skeletal dysplasias were curated from the literature and displayed in Figure 3.6c-d with residue numbering according to FGFR3c. Mutated positions occurring in both cancer and dysplasias are highlighted and labelled and include mutations at the two hotspots N540 (most commonly mutated to lysine) and K650 (most commonly mutated to glutamic acid or methionine), similar to the most frequently observed mutations in cancer. These two residues are respectively located in the FGFR KD in the αC-β4 loop (also in the 'molecular brake') and the A-loop. Other mutated residues common between cancer and dysplasias occur at I538 (also located in the αC-β4 loop adjacent to the 'molecular brake') and R669 (located in the αEF-β12 loop adjacent to the A-loop). The G697 cancer mutation hotspot (located in the αF-αG loop) has not been found mutated in dysplasias in any FGFR isoform.

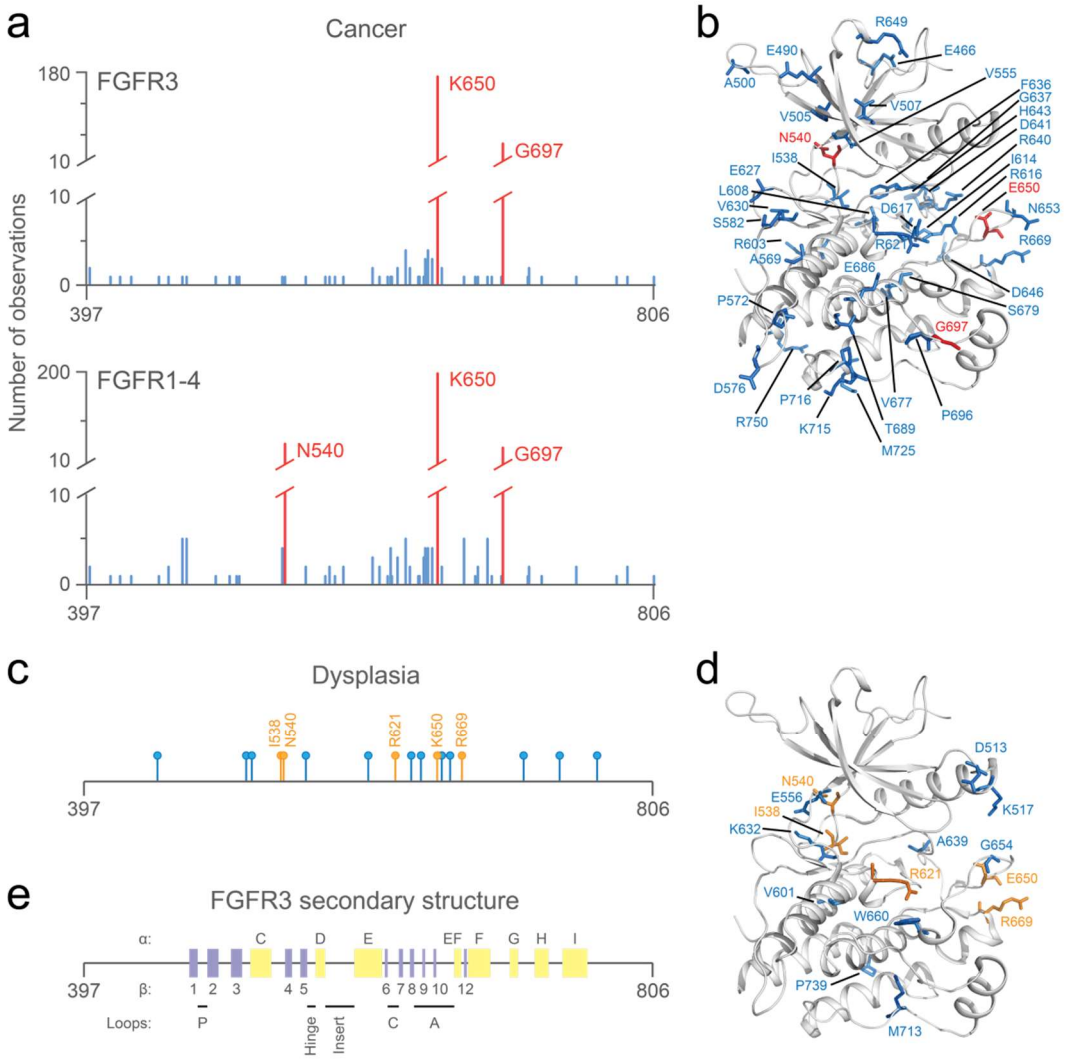


Figure 3.6 Compilation of cytoplasmic FGFR3 cancer-associated and dysplasia mutations and corresponding mutations in other FGFRs. (a) Histogram showing the total number of observations of somatic missense substitutions in the cytoplasmic region (397-806) of FGFR3 (upper panel) and corresponding mutations in FGFR1-4 (lower panel) found in cancer, from COSMIC v71 (November 2014). Mutation hotspot positions are highlighted in red and labelled with the FGFR3 residue number. See Table 3.1 for numbers of observations of mutations at each position in each FGFR. **(b)** Structural locations of the residue positions of missense substitutions in **a**, mapped on a cartoon representation of the only available FGFR3 KD crystal structure (FGFR3-K650E) [PDB: 4K33] (Huang et al., 2013b) (coloured grey). Residues at which mutations occur are found all over the KD (shown as sticks, coloured blue and labelled). The mutation hotspot positions when considering mutations in FGFR1-4 are shown as sticks, highlighted in red and labelled (note that K650 is E650 in this crystal structure). **(c)** Illustration of the positions of germline mutations found most commonly in FGFR1-4 in skeletal dysplasias curated from the literature and displayed with residue numbering according to FGFR3. Positions at which mutations have been observed in both cancer and dysplasias in FGFR3 are highlighted in orange and labelled showing five common positions mutated in both types of pathologies. **(d)** Structural locations of the mutations in **c**, mapped on a cartoon representation of the same crystal structure as in **b**. Residues at which mutations occur (fewer than those found in cancer) are shown as sticks, coloured blue and labelled. Positions found mutated in both cancer and dysplasias are shown as sticks, coloured orange and labelled. **(e)** Secondary structure elements and features of the FGFR3 cytoplasmic region (including the KD) shown to the same horizontal scale as the panels in **a** and **c**. The eight α-helices (αC-E, αEF, αF-I) and the eleven β-strands (β1-10, β12) are shown as pale yellow and lilac rectangles respectively. The key loop regions are also labelled beneath the illustration.

Table 3.1 Missense substitutions in the cytoplasmic region (397-806) of FGFR3 (and corresponding mutations in FGFR1, FGFR2 and FGFR4) in COSMIC v71 (November 2014) to complement data in Figure 3.6. The numbers of observations in cancer are given in square brackets for the total number of mutations observed at each position.

Structural location	FGFR3		FGFR1	FGFR2	FGFR4
Juxta-membrane	R399C/H	[2]	K400	K401	Q393
	I414V	[1]	L417	L418	K408
	R421Q	[1]	R425	R426	R416
	A429V	[1]	A435	S436	S424
	P449S	[1]	P455	P458	A444
	E456K	[1]	E462	E465	S451G [1]
	E466K	[1]	E472	E475K/G [4]	E461
	R469Q	[1]	R475W/Q [3]	R478	R464W [1]
β2	E490G	[1]	E496	E499	E485
β2-β3 loop	A500T	[1]	N506	K509	D495
β3	V505I	[1]	V511	V514	V500
	V507M	[1]	V513	V516	V502
αC-β4 loop	I538F	[1]	I544	I547V/T [3]	I533
	N540S	[1]	N546D/S/K [18]	N549H/K/S [35]	N535
β5	V555M	[1]	V561	V564	V550L [1]
αD	A569V	[1]	A575	A578	A564
Kinase insert	P572A	[1]	P578	P581Q [1]	P567
	D576N	[1]	E582	E585	D571
	C582F	[1]	S588T [1]	N591	P577
αE	R603Q	[2]	R609	R612T [1]	R598
	L608M	[1]	L614	L617F [1]	L603
β6	I614N	[1]	I620	I623	I609
C-loop	R616G	[1]	R622	R625Q [2]	R611Q [1]
	D617G	[1]	D623	D626	D612
	R621H	[2]	R627	R630	R616G [1]
β7-β8 loop	E627K/G/V/D	[4]	E633	E636K [1]	E622
A-loop	V630M/A	[2]	V636	V639	V625
	F636L	[1]	F642	F645	F631
	G637W	[1]	G643	G646	G632
	R640W	[1]	R646	R649	R635S [2]
	D641N/G	[3]	D647N [1]	D650	G636
	H643D/R	[4]	H649	N652	H638
	D646N/Y/G	[3]	D652	D655G [1]	D641
	K650E/M/N/Q/T	[173]	K656E/M [13]	K659E/M/N [11]	K645E [1]
	N653H/S	[2]	N659	N662	N648
αEF-β12 loop	R669Q	[1]	R675Q [1]	R678G/S [2]	R664L [1]
αF	V677I	[1]	V683	V686	V672
	S679F	[1]	S685	S688F [1]	S674
	E686K	[1]	E692	E695K [1]	E681K [3]
αF-αG loop	T689M	[1]	T695	T698	T684
	P696L	[1]	P702	P705	P691
	G697C	[44]	G703S [2]	G706	G692
αG-αH loop	K715M	[1]	K721	K724	R710
	P716H	[2]	P722	P725	P711
αH	M725I	[1]	M731	M734	G720
αI	R750C	[1]	R756	R759Q [1]	K745
C-tail	S779R	[1]	S785	S788	S773
	S787F	[1]	S794	S796F [1]	S780
	T806M	[1]	R821	T821	T802

The panel of 25 mutations analysed in this work was finally selected using a semi-biased selection process. Hotspot mutations (e.g. K650E) and mutations occurring at the same residue across more than one FGFR isoform (e.g. R669G) were selected, as it was hypothesised that properties of a mutation such as 'conservation' and/or frequent occurrence across FGFR isoforms in cancer might imply that the mutation plays a role in pathogenesis. Variants were also selected in functionally important regions of the kinase such as the A-loop (D641N) and active site (D617G) or if they co-occurred in dysplasias (e.g. I538V). Two 'resistance' mutations were selected from the literature due to their occurrence as secondary acquired mutations in pre-clinical models of resistance to targeted FGFR therapies (V555M, Y647C) (Chell et al., 2013, Crystal et al., 2014). Finally, some rarely observed mutations were selected at random, many of which have only been observed in one FGFR isoform (e.g. A500T). The final panel of representative mutations taken forward for experimental evaluation included half of all residues found to be mutated in the FGFR3 cytoplasmic region in cancer. Details of the mutations in this compiled panel are provided in Table 3.2.

3.3.2 Construction of a panel of FGFR3 mutant KD proteins

A human FGFR3 (455-768) KD construct was cloned by Tom Bunney in the pOPINS bacterial expression vector (OPPF) and another was purchased in the pJ821 bacterial expression vector (DNA2.0). The pOPINS construct comprised the native *FGFR3* cDNA sequence while the pJ821 construct was codon optimised for *E. coli* expression. Both constructs were otherwise identical in sequence boundaries and overall design. They contained an N-terminal Ulp1 protease-cleavable 6His-SUMO purification tag and solubilising '2C' mutations C482A and C582S (Mohammadi et al., 1996b) which also increase soluble expression of this construct while maintaining WT kinase activity (data not shown). These FGFR3 'WT' constructs will therefore be referred to as FGFR3-2C or 2C throughout this work. The modular design of this basic expression construct is shown in Figure 3.7a. Point mutations were introduced into this backbone by site-directed mutagenesis (SDM) using the primers in Tables 2.1 and 2.2. Constructs were also checked by sequencing to confirm the successful incorporation of mutations (for full methods, see sections 2.1.5 and 2.2.1).

Table 3.2 FGFR3 mutations selected for experimental analysis in this work. Information about cancer tissue type was extracted from COSMIC. Hotspot FGFR mutations are highlighted in red.

Structural location	FGFR3 residue	Substitution in FGFR3	Corresponding substitution			Main cancer tissue type	Dysplasia mutation?	Selected for this study	Yield ($mg/L_{E.\ coli}$)
			FGFR1	FGFR2	FGFR4				
Juxtamembrane	E466	K		K,G		CNS (Glioma)		E466K	0.13
β2-β3 loop	A500	T				Pancreas		A500T	2.85
αC-β4 loop	I538	F		V/T		Haematopoietic & lymphoid	FGFR2, 3	I538F,V	0.70, 1.00
	N540	S	D,S,K	H,K,S		Urinary tract	FGFR3	N540S,K	0.98, 0.15
β5	V555	M			L	Haematopoietic & lymphoid		V555M*	1.10
Kinase insert	P572	A		Q		CNS		P572A	2.96
	C582	F	T			Ovary		C582F	3.12
C-loop	D617	G				Upper aerodigestive tract		D617G	3.84
β7-β8 loop	E627	D,G,K,V		K		Urinary tract		E627D	2.32
A-loop	V630	A,M				Upper aerodigestive tract		V630M	2.16
	G637	W				Kidney		G637W	1.10
	D641	G,N	N			Urinary tract		D641G,N	1.69, 1.58
	H643	D,R				Urinary tract, pancreas		H643D	4.29
	D646	G,N,Y		G		Large intestine, urinary tract		D646Y	0.20
	Y647	C				Lung		Y647C*	0.39
	K650	E,M,N,Q,T	E,M	E,M,N	E	Urinary tract, skin, etc	FGFR2, 3	K650E,N	2.01, 0.76
	N653	H,S				Urinary tract		N653H	1.72
αEF-β12 loop	R669	Q	Q	G,S	L	Large intestine	FGFR2, 3	R669G,Q	0.50, 1.44
αF	V677	I				Large intestine		V677I	0.80
	G697	C	S			Upper aerodigestive tract		G697C	0.53

* Secondary acquired resistance mutation obtained from the literature (Chell et al., 2013, Crystal et al., 2014). The V555M mutation is also listed in COSMIC.

This plasmid library of FGFR3-2C and mutant KDs was used to express and purify mutant KD proteins from *E. coli* C41 (DE3) λ37 cells pre-transformed with an accessory plasmid pCDFDuet-1™ (Novagen) containing genes for λ-phosphatase and the CDC37 co-chaperone. The λ-phosphatase was included to remove traces of phosphorylation from the KDs during expression (since the KDs are capable of autophosphorylation) while FGFR3 is a client of the CDC37 co-chaperone which aids in folding of FGFR3 and improves its solubility (Laederich et al., 2011). Additionally, in the absence of the λ-phosphatase, KD yields were lower by several fold (data not shown) so inclusion of the λ-phosphatase was crucial for expressing reasonable yields of soluble KD proteins. The expression and purification workflow is summarised in Figure 3.7a-b and full methods are detailed in section 2.6.1. Yields of different mutants varied greatly, with some being much easier to produce and others producing just a few micrograms of purified KD per litre of *E. coli* culture (Table 3.2). Mutations I538F, I538V, V555M and G637W were last to be added to the panel and following generation of the mutant plasmid constructs, these KD proteins were purified and provided by Tom Bunney. Peak fractions from the final step size exclusion chromatography (SEC) that were found to contain monomeric proteins (Figure 3.7c) were concentrated and used in further experiments. The final purity of FGFR3 mutant KD proteins used in this work was estimated to be over 95% from SDS-PAGE analysis (Figure 3.7d). Although a lot of care was taken during expression and purification (as described above) to ensure the KDs prepared were unphosphorylated, all purified KD proteins were checked by native mass spectrometry (see Figure 3.7e for example mass spectra) and native PAGE (data not shown) to confirm their zero-phosphorylation status prior to use in experiments.

3.3.3 Development of an *in vitro* kinase activity assay

The first aim of this work was to directly and comparatively screen the comprehensive panel of FGFR3 variant KD proteins to differentiate potential oncogenic drivers from passenger mutations. This required a robust assay platform to perform screening, activity and kinetics assays in a relatively small scale and medium-throughput format. The kinase activity assay was developed using FGFR3-2C and using the ADP-Glo™ kinase assay (Promega). Figure 3.8 schematically illustrates the rationale of the ADP-Glo™ methodology. Briefly, this assay produces a luminescent output signal that is representative of and proportional to the amount of ADP produced during an ATPase (or kinase) reaction. Assays were performed with and without a poly-Glu-Tyr ($polyE_4Y_1$) peptide (Sigma) which has been shown to be a suitable substrate in pilot experiments where it was phosphorylated by FGFR KDs (data not shown). The assay was designed and optimised for a 96-well plate format to allow medium-throughput scale manual

experiments to be performed. All experiments were performed in this format unless otherwise indicated.

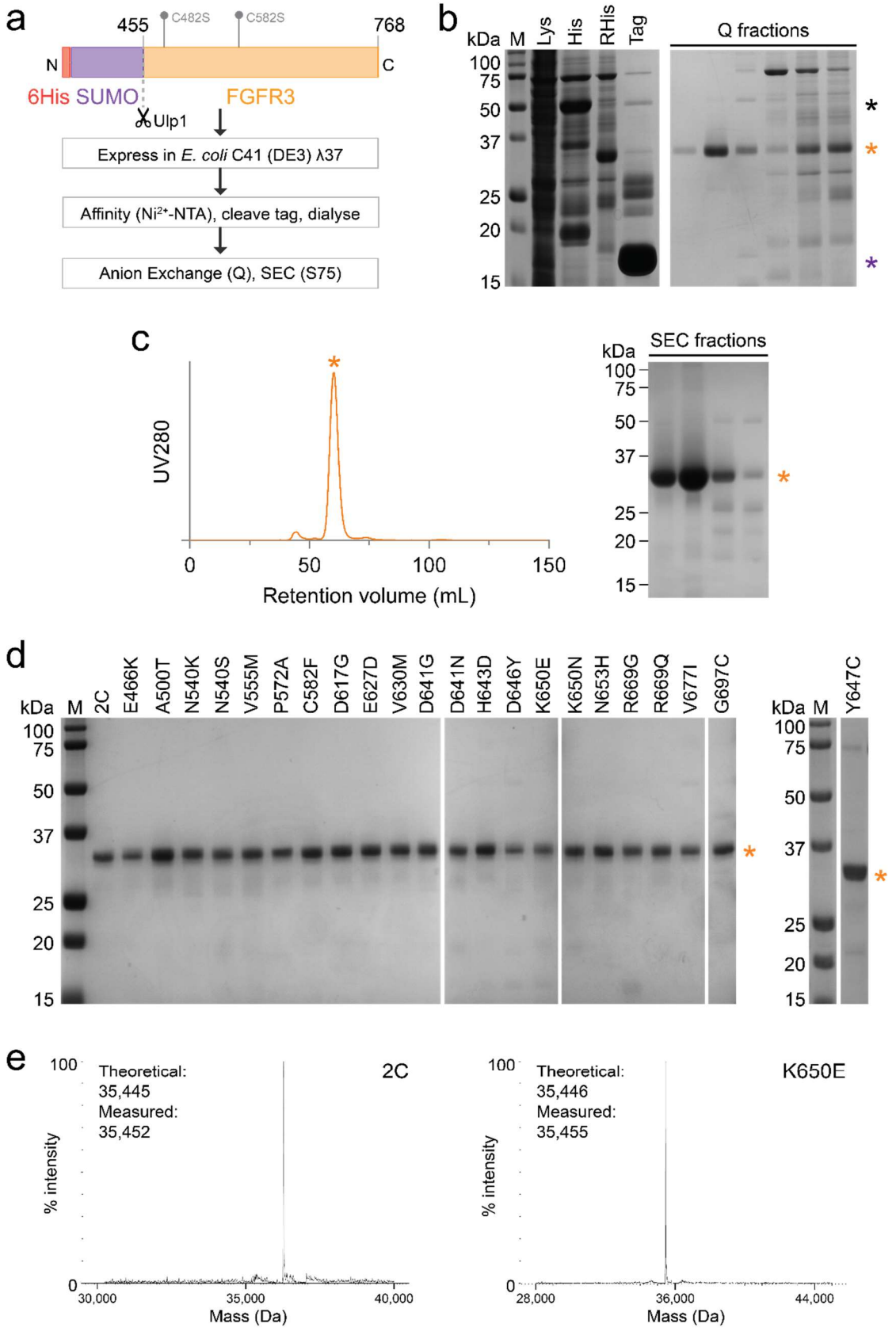


Figure 3.7 FGFR3 KD construct design, expression, purification and quality control. (a) Construct design, expression and purification workflow used for the production of FGFR3-2C and variant KD proteins. The FGFR3 (455-768) construct (in either the pOPINS or pJ821 vector backbone) harboured two solubilising cysteine point mutations (C482A and C582S, annotated) and the relevant pathogenic point mutations (Table 3.2). It comprised an N-terminal 6His-SUMO purification tag cleavable by Ulp1 protease (indicated by scissors). This construct was transformed into *E. coli* and subjected to the depicted expression and purification workflow. SEC – size exclusion chromatography, S75 – Superdex 75. **(b)** Representative Coomassie stained SDS-PAGE gels for all FGFR3 variants demonstrating the purification workflow. Protein is shown in the clarified *E. coli* soluble lysate (Lys), elution following Ni^{2+}-NTA chromatography (His), flowthrough (RHis) and elution (Tag) following Ulp1 protease cleavage of the 6His-SUMO tag and 'reverse' Ni^{2+}-NTA chromatography, and fractions from 'Q' anion exchange chromatography. 6His-SUMO FGFR3 KD protein is marked by a black asterisk, Ulp1-cleaved FGFR3 KD protein (tag-free) is marked by an orange asterisk, and the cleaved 6His-SUMO tag is marked by a purple asterisk. Protein from the first three fractions shown from the Q chromatography were pooled and concentrated prior to SEC. **(c)** Representative SEC chromatogram (left panel) and Coomassie stained SDS-PAGE gel (right panel) for all FGFR3 variant KD proteins showing that all proteins eluted in a single peak comprising monomeric protein in the final chromatographic purification step (marked by a single asterisk). All four SEC fractions shown were pooled, concentrated and used in further experiments (and similarly for all other variants). **(d)** Coomassie stained SDS-PAGE gels for FGFR3 variant KDs produced for this work (I538F, I538V and G637W prepared by Tom Bunney are not shown) showing that the purity of the proteins following the final SEC purification step was over 95% for all proteins (left panels). Y647C is shown separately (right panels) for a sample following the anion exchange purification step (prior to SEC) as an SDS-PAGE gel is not available for this mutant following SEC. Bands corresponding to FGFR3 KD proteins in both panels are marked by a single asterisk. **(e)** Native LC-MS spectra of purified KD proteins following SEC and storage at -80°C. Only spectra for FGFR3-2C (left panel) and FGFR3-K650E (right panel) are shown as examples, indicating the theoretical calculated masses and LC-MS measured masses in Daltons (Da). Although the spectra shown do not span the expected dimeric mass of the KD proteins, no peaks were observed for the FGFR3 KDs at this mass (approximately 73,000 Da). For full experimental details see sections 2.7 and 2.8.

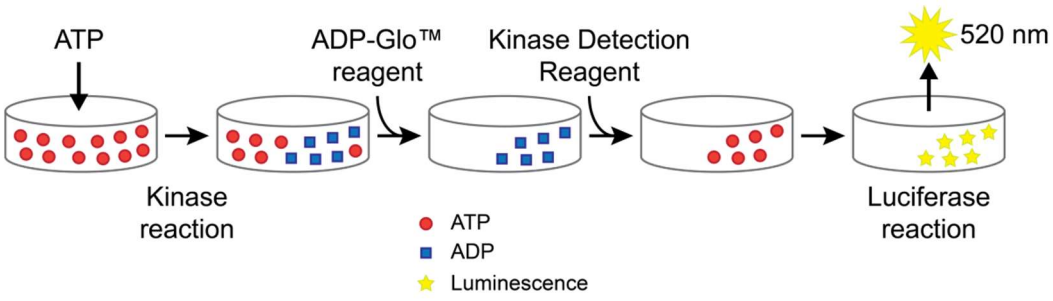
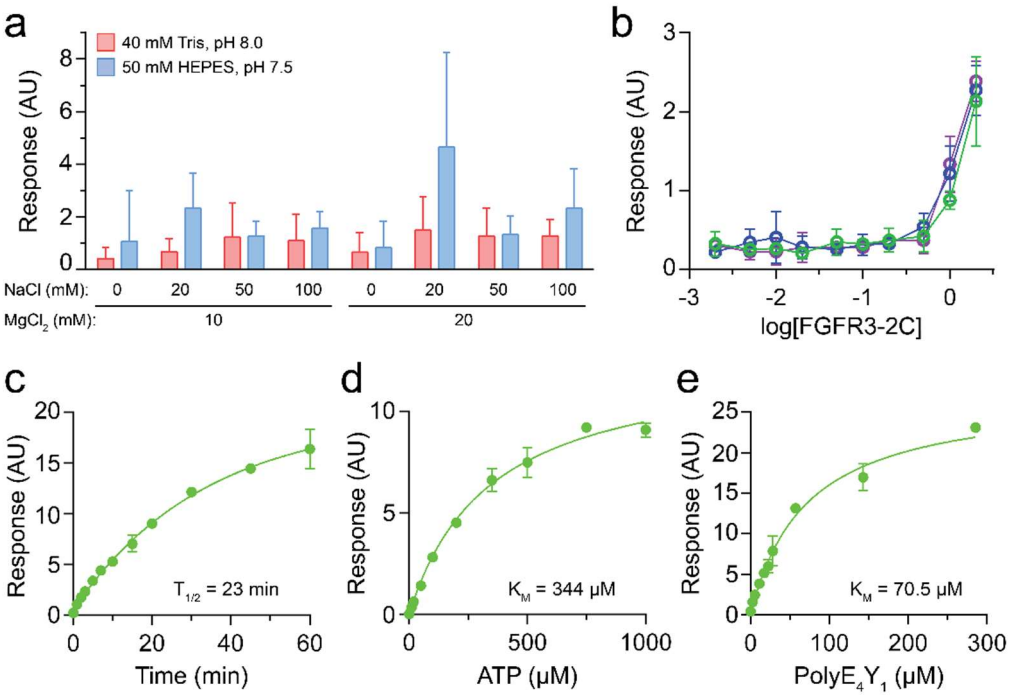


Figure 3.8 Rationale behind the ADP-Glo™ methodology. Illustration of the rationale behind the ADP-Glo™ methodology (Promega). Kinase reactions were started at time = 0 seconds by the addition of Ultra Pure ATP (Promega) or kinase, depending on the experiment type. During the kinase reaction, ATP was converted to ADP by the KDs. Reactions were stopped by the addition of ADP-Glo™ reagent into the kinase reaction. This reagent also depleted remaining ATP in the reaction over the 40 minute incubation period. Next, Kinase Detection Reagent was added to the reactions to convert ADP produced by the kinase reaction into ATP over a 30-60 minute incubation period, depending on the starting concentration of ATP. This ATP then fuelled a luciferase reaction, producing a luminescent signal that was detected using a plate reader configured to measure luminescence at 520 nm. For full experimental details see section 2.9.



Assay optimisations were initially performed for constituents of the kinase reaction buffer (KRB), reaction time, ATP concentration and polyE$_4$Y$_1$ concentration (Figure 3.9) using FGFR3-2C to ensure that signal linearity of all parameters was maintained during the course of the assays. The optimal composition of KRB (used for all subsequent experiments) was determined to comprise 50 mM HEPES buffer (pH 7.5), 20 mM NaCl and 20 mM $MgCl_2$ based on the buffer screening experiment shown in Figure 3.9a. Also present in all experiments was 100 µM of both phosphatase inhibitor sodium orthovanadate and reducing agent TCEP to prevent dephosphorylation and inhibit aggregation respectively during the reactions. The reaction time for further experiments was selected to be 10 minutes, below the 'half-time' $T_{1/2}$ (the time at which half of the ligand binding sites become occupied with ligand) calculated here for 2C (23 minutes, Figure 3.9c) to ensure strict signal linearity while maintaining signal intensity. Similarly for concentrations of ATP and polyE$_4$Y$_1$, the affinity constant (K_M) for each substrate determined here (344 µM and 70.5 µM respectively, Figure 3.9d-e) guided the concentrations of each substrate used in further experiments. These concentrations varied for different types of assays and the exact concentrations used in specific experiments are detailed in the Figure legends for the relevant experiments.

Figure 3.9 Optimisation of the kinase activity assay using FGFR3-2C. (a) Screening of kinase reaction buffer (KRB) components including varying the pH (using indicated buffers), NaCl and $MgCl_2$ concentrations as indicated. These experiments were uniquely performed using a combination of manual and semi-automated liquid dispensing into 384-well plates for kinase reactions with a total volume of 4 µL per reaction and subsequently similarly scaled volumes of ADP-Glo™ reagents. The results shown are for 1 nM FGFR3-2C, 0-100 mM NaCl and 10-20 mM $MgCl_2$, representative of two similar independent experiments performed additionally with other variable constituents (which did not alter the trend shown in the graph and so are not shown): FGFR3-2C (1, 10, 100 and 1000 nM), bovine serum albumin (0 and 100 µg/mL), four detergent conditions (0% or 0.005% Brij-35, CHAPSO and Tween-20), four concentrations of ATP (0,10,100 and 1000 µM) and $polyE_4Y_1$ peptide (0.0286, 0.286, 2.86 and 28.6 µM). All reactions also contained 100 µM sodium vanadate (phosphatase inhibitor) and 100 µM TCEP (reducing agent). Each condition was tested in at least triplicate reactions per assay, and error bars represent the standard deviation of six repeats for each condition. The optimal conditions from this experiment (50 mM HEPES pH 7.5, 20 mM NaCl, 20 mM $MgCl_2$, 100 µM sodium vanadate and 100 µM TCEP) formed the basic composition of KRB used for experiments in this work. AU – arbitrary units. **(b)** Autophosphorylation experiments performed using increasing concentrations of FGFR3-2C between 0-2 µM, 50 µM ATP and 45 minute kinase reactions, displayed as log dose-response curves for three independent experiments (green, blue and purple) of at least five independent experiments performed under identical conditions and using at least two independent 2C protein preparations. Each data point was measured in triplicate and error bars represent the resulting standard deviation for each concentration of kinase. The curves connecting the data points are not fits to the data and are only shown for visual clarity. **(c)** Time course experiment performed using 0.5 µM FGFR3-2C, 429 µM $polyE_4Y_1$, 500 µM ATP, and time points between 0-60 minutes. Each data point was measured in triplicate and error bars represent the resulting standard deviation at each time point. Data were fitted to a one phase exponential model for pseudo-first order association kinetics (equation **3**) and used to calculate the 'half-time' ($T_{1/2}$, indicated) for the reaction. **(d)** ATP titration experiment performed using 0.5 µM FGFR3-2C, 429 µM $polyE_4Y_1$, 0-1000 µM ATP and 45 minute reactions. The graph represents results from two independent experiments performed under identical conditions. Each data point was measured in triplicate in each assay and error bars represent the resulting standard deviation for each concentration of ATP. Data were fitted to the Michaelis-Menten equation (equation **4**), and used to calculate the affinity constant for ATP (K_M, indicated) for the reaction. **(e)** Substrate titration experiment performed using 0.5 µM FGFR3-2C, 0-286 µM $polyE_4Y_1$, 500 µM ATP and 45 minute reactions. The graph represents results from six independent experiments performed under similar conditions. Each data point was measured in triplicate in each assay and error bars represent the resulting standard deviation for each concentration of $polyE_4Y_1$. Data were fitted to the Michaelis-Menten equation (equation **4**), and used to calculate the affinity constant for $polyE_4Y_1$ (K_M, indicated) for the reaction. For full experimental details for **b**-**e** see section 2.9.

3.3.4 Autophosphorylation screens reveal potential oncogenic drivers

The optimised kinase activity assay was then used to screen the panel of twenty six purified FGFR3 KD proteins for kinase hyperactivity. Purified FGFR3-2C and variant KDs were first incubated at 21-22°C with ATP in the presence of 20 mM Mg^{2+} in a set of experiments designed to assess the impact of FGFR3 KD mutations on autophosphorylation compared to FGFR3-2C. Output luminescence was measured at a fixed time point over a range of kinase concentrations (Figure 3.10a) or at a single kinase concentration (Figure 3.10b). All dose response and single kinase concentration experiments were performed in triplicate, with at least two or three independent dose response experiments performed with each variant. Subsequently, assays were also performed with two independent preparations of protein for several (randomly selected)

mutants and with four independent preparations of WT protein, to eliminate any possibilities of batch-to-batch variation. No such variation was seen, as demonstrated using FGFR3-2C as an example (Figure 3.9b). Both types of autophosphorylation assays produced comparable, reproducible results which will be reviewed in parallel.

Eleven of the 25 variants exhibited greater than 3-fold activity compared to FGFR3-2C. These are found at eight distinct residue positions on the KD, all of which are conserved across the four FGFRs (the only exception is FGFR3-D641 which is a glycine in FGFR4). The two hotspot mutations N540K and K650E, which are found most frequently replaced at each residue position across FGFRs, caused some of the largest increases in autophosphorylation of around 40- and 45-fold respectively. The mutations N540S and K650N which are less frequently observed at the same residue positions were also less hyperactive, although they were still activating by around 10- and 18-fold respectively. Mutations at position K650 have been thoroughly studied in the context of congenital diseases and cancer, and shown to be activating in FGFR3 (see section 3.1.5). This mutation therefore served as a good positive control for all experiments performed with mutations in this work. The mechanism of activation by K650E has also been elucidated (see section 3.1.5); briefly, this mutation stabilises the active KD conformation, independent of tyrosine phosphorylation and FGF ligand stimulation, by mimicking the action of phosphorylated Y647 in the A-loop (Figure 3.4). Mutations at the corresponding residues for FGFR3 N540 and K650 in other FGFRs have also been shown to be activating *in vitro* (Chen et al., 2007b, Lew et al., 2009, Byron et al., 2013, Chen et al., 2013b). This work shows that they are also activating in FGFR3.

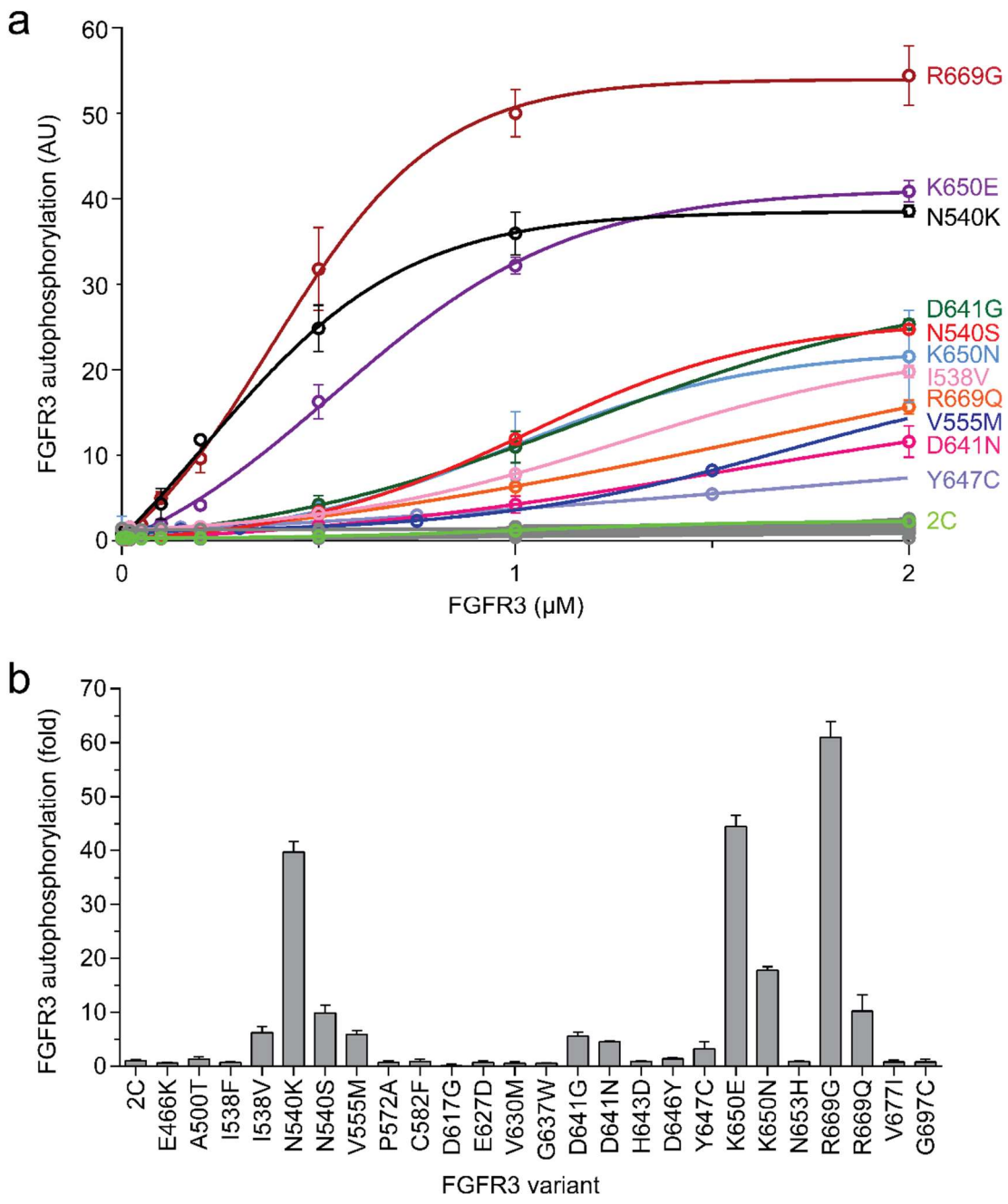


Figure 3.10 Autophosphorylation experiments. (a) Autophosphorylation experiments performed using increasing concentrations of FGFR3 variants between 0-2 μM, 50 μM ATP and 45 minute kinase reactions, displayed as dose response curves fitted with a saturation dose response binding equation for visual clarity (equation **1**). Data shown for FGFR3-2C are representative of six identical independent experiments (using at least three independent protein preparations), while data for other variants are representative of at least two identical independent experiments each (using two independent protein preparations for many variants). Each data point was measured in triplicate in each assay and error bars represent the resulting standard deviation for each concentration of kinase. Eleven mutations found to be more activating than 2C are highlighted in different colours (used throughout this work) and labelled. Mutations found to not be more activating than 2C are coloured grey and are not labelled. AU – arbitrary units. **(b)** Autophosphorylation screening experiments performed using 1 μM FGFR3 variants, 50 μM ATP and 45 minute kinase reactions displayed as a histogram showing fold change compared to 2C (normalised 2C = 1). Each data point was measured in triplicate and error bars represent the resulting standard deviation for each variant. Some variants were also included in a second independent assay under identical conditions. For full experimental details see section 2.9.

Interestingly, the mutation displaying the highest activation in both of these assays was R669G, a non-hotspot mutation in the αEF-β12 loop displaying an autophosphorylation increase of more than 60-fold compared to FGFR3-2C. The other variant found at the same residue position (R669Q) was less hyperactive, but still activating by 10-fold. This work shows for the first time that pathogenic mutations at this position in FGFR3 activate the KD. The corresponding mutation to FGFR3-R669G in FGFR2 (R678G) was once assessed for its activation status in the context of skeletal dysplasias and was shown to be activating *in vitro* (Chen et al., 2007b) but was not found to be as activating compared to FGFR2-WT as in this work, and the authors made no further comments about this mutation. Additionally, replacement of the corresponding residue in FGFR1 to glycine (R675G) also results in activation of the KD compared to FGFR1-2C (assessed using a construct similar to the FGFR3-2C used in this work) (Patani et al., 2016). The small discrepancies between these results might reflect subtle structural and/or local sequence differences between the two FGFR KDs which present as small differences in kinase activation by the same mutation at the corresponding residue in different FGFR isoforms.

The other five mutations found to moderately activate kinase autophosphorylation were I538V adjacent to the 'molecular brake' and recently described as part of the 'DFG-latch', V555M in the ATP binding pocket (the FGFR3 gatekeeper residue) and A-loop mutations D641G, D641N and Y647C displaying 6-, 6-, 5.5-, 4.5- and 3-fold activation respectively compared to FGFR3-2C (Figure 3.10b). The I538V mutation is also found in hypochondroplasia (Helsten et al., 2015) and interestingly, mutation to phenylalanine at this residue (I538F) was not found to be activating. The gatekeeper mutation V555M was identified as a secondary mutation acquired as a resistance mechanism in response to pre-clinical treatment of a human tumour model cell line with an FGFR inhibitor (Chell et al., 2013), yet this mutation has been curated into the COSMIC database. Similar to the result shown here, previous *in vitro* activity assays in FGFR3 (Bunney et al., 2015), FGFR2 (Byron et al., 2013) and other kinases (Azam et al., 2008) have shown that mutations to bulkier residues at the gatekeeper position are activating. The other acquired resistance mutation Y647C, observed in response to pre-clinical treatment of a model human tumour cell line with an EGFR inhibitor (Crystal et al., 2014) is also moderately activating in this assay. Mutations at FGFR3-D641 found in the A-loop or corresponding mutations in other FGFRs have not been previously studied, so this assay showed for the first time that mutations at this position are also activating in FGFRs. Interestingly, mutations occur at the corresponding and conserved position to FGFR3-D641 in the related RTKs KIT (KIT-D816) and PDGFRα (PDGFRα-D842). They form the largest mutation hotspots in COSMIC for both RTKs with approximately 2000 and 600 observed mutations in cancer at this position in each RTK respectively. This region

('DFGLARD', 635-641 in FGFR3) is 100% conserved in both RTKs and FGFR3. The most frequently observed mutations are KIT-D816V and PDGFRα-D842V, almost exclusively observed in mastocytosis for KIT-816V (Nagata et al., 1995, Furitsu et al., 1993), or in soft tissues for PDGFRα-D842V (Heinrich et al., 2003). Both mutations have also been shown to be constitutively, ligand-independently activating, as first shown for KIT-D816V (Kitayama et al., 1995, Furitsu et al., 1993) and for PDGFRα-D842V (Heinrich et al., 2003). Although the corresponding mutation to valine at this position has not been observed in FGFRs in cancer, it is interesting that different replacements at D641 are also rarely observed in FGFRs (mostly FGFR3) in cancer (Table 3.1) where they are also activating (Figure 3.10).

Fourteen of the panel of 25 variants displayed either very little or no effect on activation with some variants even being less activating compared to FGFR3-2C (Figure 3.10). Most of these variants are only rarely observed in cancers. However, included in these non-activating mutations is the hotspot mutation G697C, located in the αF-αG loop, whose controversial activation status and relevance in cancer is currently unresolved. In this work, the mutation is not activating, supporting the newer study suggesting it is irrelevant in cancer (Aubertin et al., 2007). Remarkably, KD activity was completely abrogated by mutations D617G and G637W. These mutations occur at catalytically essential residues conserved across the kinome; they are part of the HRD and DFG motifs located in the C-loop and A-loop respectively. Mutation of the catalytic aspartate in the HRD motif (D617G) renders the KD incapable of phosphoryl transfer (Valiev et al., 2003). Mutations affecting the DFG motif, which coordinates the ATP and Mg^{2+} co-factors during phosphorylation, also render the FGFR3 KD catalytically inactive. This is observed across kinases which appear mutated in this motif in cancer more frequently than expected (Greenman et al., 2007, McSkimming et al., 2015). This is the first time kinase activity has been assessed for the G637W mutation in any kinase. Overall, the importance of cancer-associated mutations that reduce or abolish FGFR kinase activity is not yet known, but it has been suggested that inactivating mutations may have a context-dependent tumour protective role (Gartside et al., 2009, Turner and Grose, 2010). It is also possible that they are passenger mutations with no oncogenic role and are rarely detected in cancers due to the increased genomic instability of tumours.

In summary, from the autophosphorylation experiments, mutations at the three residues N540, K650 and R669 (located in the 'molecular brake', A-loop and αEF-β12 loop respectively) were found to be the most activating, with all six variants at these three positions displaying greater than 10-fold activation compared to FGFR3-2C (Figure 3.10b). Interestingly, all three mutations are also observed in dysplasias in both FGFR3

and FGFR2 (Figure 3.6c) where they are responsible for causing the most severe forms of skeletal abnormalities (Wilkie, 2005), albeit with less activating mutations at the same residues in dysplasias causing less severe pathogenic phenotypes. This work also shows that different replacements at the same residue can result in differential levels of kinase activation, as has been previously shown for mutations at K650 and N540 respectively in FGFR3 and other FGFRs (Webster et al., 1996, Bellus et al., 2000, Chen et al., 2007b, Chen et al., 2013b). Interestingly, a different replacement at the position corresponding to FGFR3-R669 in FGFR1 (R675E) has been shown to reduce kinase activity compared to FGFR1-WT (Kobashigawa et al., 2015) and it was suggested that this residue is important for *trans*-autophosphorylation of the first A-loop tyrosine (Y653 in FGFR1). Therefore, it is likely that different substitutions at any residue might differently affect kinase activation. These subtle effects of mutations on kinase activation can only be investigated, and the mechanisms deciphered, through experimental testing.

3.3.5 Substrate phosphorylation assays confirm activation status of mutations

The panel of FGFR3 KD proteins was subsequently screened in two similar and complementary experiments to the autophosphorylation screens described above. This time (and in all subsequent assays), each reaction also contained the polyE_4Y_1 peptide substrate (also simply referred to as 'substrate' in this work) in order to assess the impact of FGFR3 KD mutations on substrate phosphorylation compared to FGFR3-2C. As with the autophosphorylation experiments, kinase dose response assays and single concentration experiments were set up (see section 2.9 for more experimental detail). The results were mostly consistent with data from autophosphorylation experiments so for clarity, data have been presented grouped according to most activating (including moderately activating mutations at the same positions), moderately activating and non-activating mutations based on the autophosphorylation screens (Figure 3.11). In the context of substrate phosphorylation the most activating mutations were still N540K, K650E and R669G (Figure 3.11a). Closely following were the other variants at these residues (N540S, K650N and R669Q) as well as the moderately activating mutations (I538V, V555M, D641G and D641N) (Figure 3.11b). The only moderately activating mutation from autophosphorylation experiments that was not activating in the context of substrate phosphorylation was Y647C (Figure 3.11b). This is partly consistent with the autophosphorylation screen (Figure 3.10b) where this mutation was also the least activating of all activating mutations (3-fold activation). Strikingly, the complete abolishment of kinase activity could be seen most clearly for the HRD and DFG motif mutations D617G and G637W in substrate phosphorylation assays (Figure 3.11c), showing that these 'kinase dead' mutations render the kinase catalytically inactive.

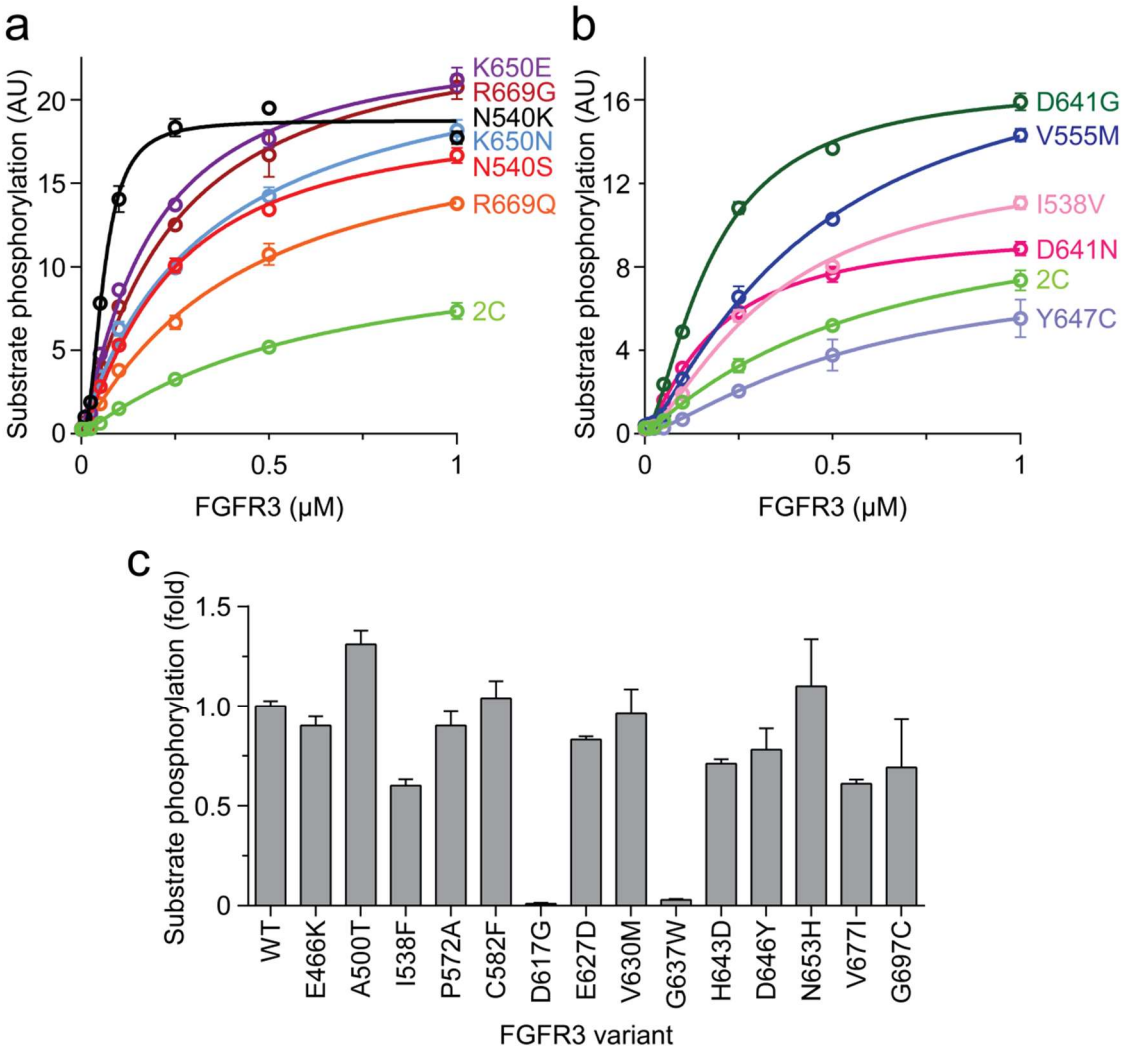


Figure 3.11 Substrate phosphorylation experiments. Substrate phosphorylation experiments performed using increasing concentrations of FGFR3 variants between 0-1 μM, 28.6 μM polyE$_4$Y$_1$, 500 μM ATP and 10 minute kinase reactions for FGFR3-2C and activating variants at positions N540, K650 and R669 **(a)** and FGFR3-2C and activating variants at positions I538, V555, D641 and Y647 **(b)**, displayed as dose response curves fitted with a saturation dose response binding equation for visual clarity (equation **2**). Data are representative of two independent experiments each for most variants under identical conditions. Each data point was measured in triplicate in each assay and error bars represent the resulting standard deviation for each concentration of kinase. Colouring of variants is the same as in Figure 3.10. AU – arbitrary units. **(c)** Substrate phosphorylation screening experiments performed using 1 μM FGFR3 variants, 14.3 μM polyE$_4$Y$_1$, 50 μM ATP and 45 minute kinase reactions displayed as a histogram showing fold change compared to 2C (normalised 2C = 1). Each data point was measured in triplicate and error bars represent the resulting standard deviation for each variant. Some variants were included in a second independent assay under identical conditions. For full experimental details see section 2.9.

These direct quantitative and comparative measurements of kinase activity *in vitro* demonstrate that different pathogenic FGFR mutations are capable of impacting and modulating kinase activation (and inactivation) to different degrees. The results suggest that while some highly activating mutations are found at genetically-defined hotspots, some are not and such non-hotspot activating mutations could be underrepresented in

mutation databases (see section 3.4). Their importance in cancer might therefore be undermined in studies that tend to focus only on the most commonly occurring mutations in cancer mutation databases. Additionally, it has been shown that hotspot mutations based solely on genetic information (such as G697C) might not necessarily be activating. Therefore, these data reinforce the importance of experimental verification of the abundant information in cancer mutation databases. Experimental assessments of the functional impact of genetically identified cancer mutations to differentiate potential oncogenic drivers from passenger mutations can then improve drug discovery and boost the success of candidate drugs in clinical trials.

3.3.6 Highly activating mutations cause partial ligand-independent activation

Screening the panel of FGFR3 variants showed that mutations impact kinase activation to differing degrees. In order to determine which mutations confer kinase activation similar to fully- activated WT, a phosphorylated active form of FGFR3-2C was prepared. Purified unphosphorylated FGFR3-2C KD was incubated with 25 mM Mg^{2+} and 10 mM ATP for 45 minutes and the resulting mixture of phosphorylated species were separated using stringent anion exchange chromatography (see section 2.7 for full experimental details). The ultimate-eluting anion exchange peak was found to contain fully phosphorylated FGFR3-2C KD ('4p', Figure 3.12a). This was concentrated and used in subsequent experiments where it is referred to as p2C. Substrate phosphorylation experiments were set up exactly as in section 3.3.5 to directly compare kinase activity of the FGFR3 KD proteins 2C, p2C and one the most highly activating variants K650E (Figure 3.12b-c). Surprisingly, it was found that even such a highly activating mutant only partially activated the KD in a ligand independent fashion compared to fully-active 'WT' p2C under these *in vitro* conditions. This is similar to the results previously observed for activating FGFR2 mutations using a different *in vitro* activity assay (Chen et al., 2013b).

The structural locations of activating mutations identified in this work through the *in vitro* activity assays are shown mapped on the FGFR3 KD in Figure 3.13a. Analysis of the primary sequences of FGFR1-4 in the regions at which the most activating mutations occur reveals near perfect sequence conservation of residues in the kinase hinge, 'molecular brake', A-loop and αEF-β12 loop. Therefore, previous mechanistic analyses of activation mechanisms for corresponding mutations in FGFR1 and FGFR2 could help to explain those in FGFR3. Activation mechanisms involving K650E and other replacements at the same residue have been thoroughly elucidated and have already been discussed in section 3.1.5 and shown in Figure 3.4. The basis for partial KD activation by different FGFR2-K659 mutations has particularly been explained to be a result of subtle structural differences that translate to small changes in the two key A-loop

and ‘molecular brake’ hydrogen bonding networks that determine the dynamic activation of the KD (Chen et al., 2013b). The other FGFR3 A-loop mutations found to be moderately activating in this work (D641G, D641N and Y647C) might also follow similar mechanisms of activation, by affecting these two key interaction networks to a lesser degree.

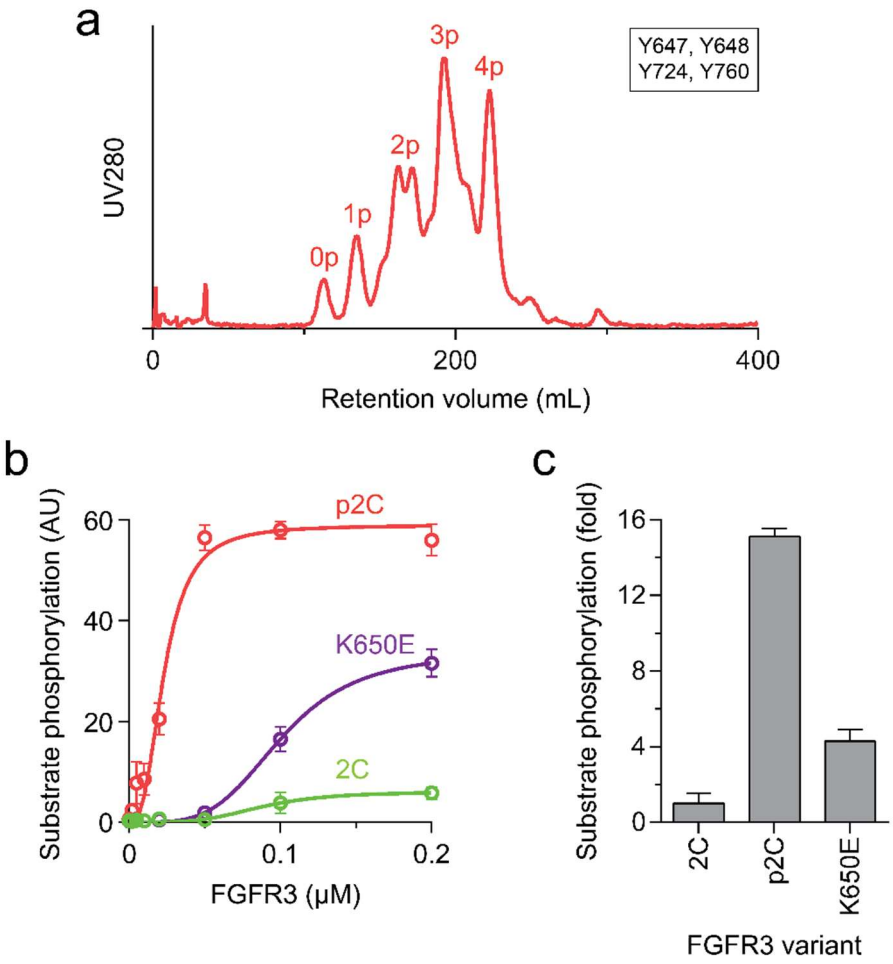


Figure 3.12 Mutations only confer partial ligand independent KD activation. (a) Purification of fully phosphorylated active FGFR3-2C (p2C). Recombinant unphosphorylated FGFR3-2C KD proteins at 2-10 mg/mL were incubated with 25 mM $MgCl_2$ and 10 mM ATP at 22°C for 45 minutes. Kinase reactions were stopped with 50 mM EDTA, desalted and loaded on to a 1 mL Resource Q anion-exchange column (GE Healthcare). Five peaks eluted, corresponding to 0p, 1p, 2p, 3p and 4p phosphorylated forms of FGFR3-2C (the four tyrosines in the KD constructs that undergo autophosphorylation are indicated in the top right corner). The 4p peak fractions contained p2C KD protein used in subsequent kinase assays. For full experimental details see section 2.7. **(b)** Substrate phosphorylation experiments performed using increasing concentrations of the indicated FGFR3 variants between 0-0.2 μM, 28.6 μM $polyE_4Y_1$, 50 μM ATP and 45 minute kinase reactions, displayed as dose response curves fitted with a saturation dose response binding equation for visual clarity (equation **2**). Each data point was measured in triplicate and error bars represent the resulting standard deviation for each concentration of kinase. AU – arbitrary units. **(c)** Substrate phosphorylation screening experiments performed using 0.1 μM FGFR3 variants, 28.6 μM $polyE_4Y_1$, 50 μM ATP and 45 minute kinase reactions displayed as a histogram showing fold change compared to 2C (normalised 2C = 1). Data are representative of two independent experiments performed using two different concentrations of kinase. Each data point was measured in triplicate in each assay and error bars represent the resulting standard deviation for each variant. For full experimental details see section 2.9.

In the kinase hinge region, activation mechanisms involving mutation of the gatekeeper residue (FGFR3-V555) have already been discussed based on previous studies of the corresponding mutation (V561M) in FGFR1 (see section 3.1.6 and Figure 3.5). For mutations at N540 in the kinase hinge, activation mechanisms are proposed based on previous analyses of the corresponding mutations in FGFR1 and FGFR2 (Chen et al., 2007b, Lew et al., 2009). The first study shows that FGFR2-N549 mutations cause direct disengagement of the molecular brake, to differing degrees depending on the amino acid substitution, directly relieving KD autoinhibition. This mechanism is likely to apply to FGFR3 as well (Figure 3.13b) due to 100% conservation of the molecular brake residues in all FGFR KDs. The second study shows that FGFR1-N546K disrupts the order of sequential tyrosine autophosphorylation compared to FGFR1-WT, allowing a higher level of basal KD activation without the need for A-loop phosphorylation. It is unclear whether this might also apply to FGFR3, which has two fewer tyrosines in its KD than FGFR1 and FGFR2, and whose sequential order of autophosphorylation has not yet been investigated. One other activating mutation in the kinase hinge (I538V) occurs at a position that normally contributes to the 'molecular brake' interactions (Figure 3.13b) and has also been recently described to form part of the 'DFG-latch' which contributes to KD activation through forming an allosteric connection that relays information between the active A-loop and the 'molecular brake' (Chen et al., 2017). Therefore, mutation of this isoleucine to valine might activate the KD by abrogating some autoinhibitory interactions in the 'molecular brake' and through allosteric mechanisms involving local structural changes in the 'DFG-latch'.

The mechanism of activation of the R669G mutation, which resulted in the largest increase in kinase activation of all mutations tested in this work, had not been previously assessed in any FGFR. Therefore, the crystal structure of FGFR1-R675G (the corresponding mutation which is also activating in FGFR1) was solved by Nethaji Thiyagarajan, to begin to understand the activation mechanism (Patani et al., 2016). Comparison of this structure with those of inactive and active FGFR1-WT confirmed that this mutation causes the KD to adopt all aspects of the typical FGFR active conformation (Figure 3.13c-e), shifting conformational equilibrium towards the active state by an allosteric mechanism similar to that proposed for other activating FGFR mutations (Chen et al., 2007b, Huang et al., 2013b). Additionally, while the autoinhibitory interactions at the 'molecular brake' are lost in this mutation, they are not abrogated to the same degree as fully phosphorylated and activated FGFR1-WT (Figure 3.13f). These subtle structural differences might partly explain why even this most activating mutation only results in partial ligand independent KD activation. Overall, KD activation is a dynamic and graded

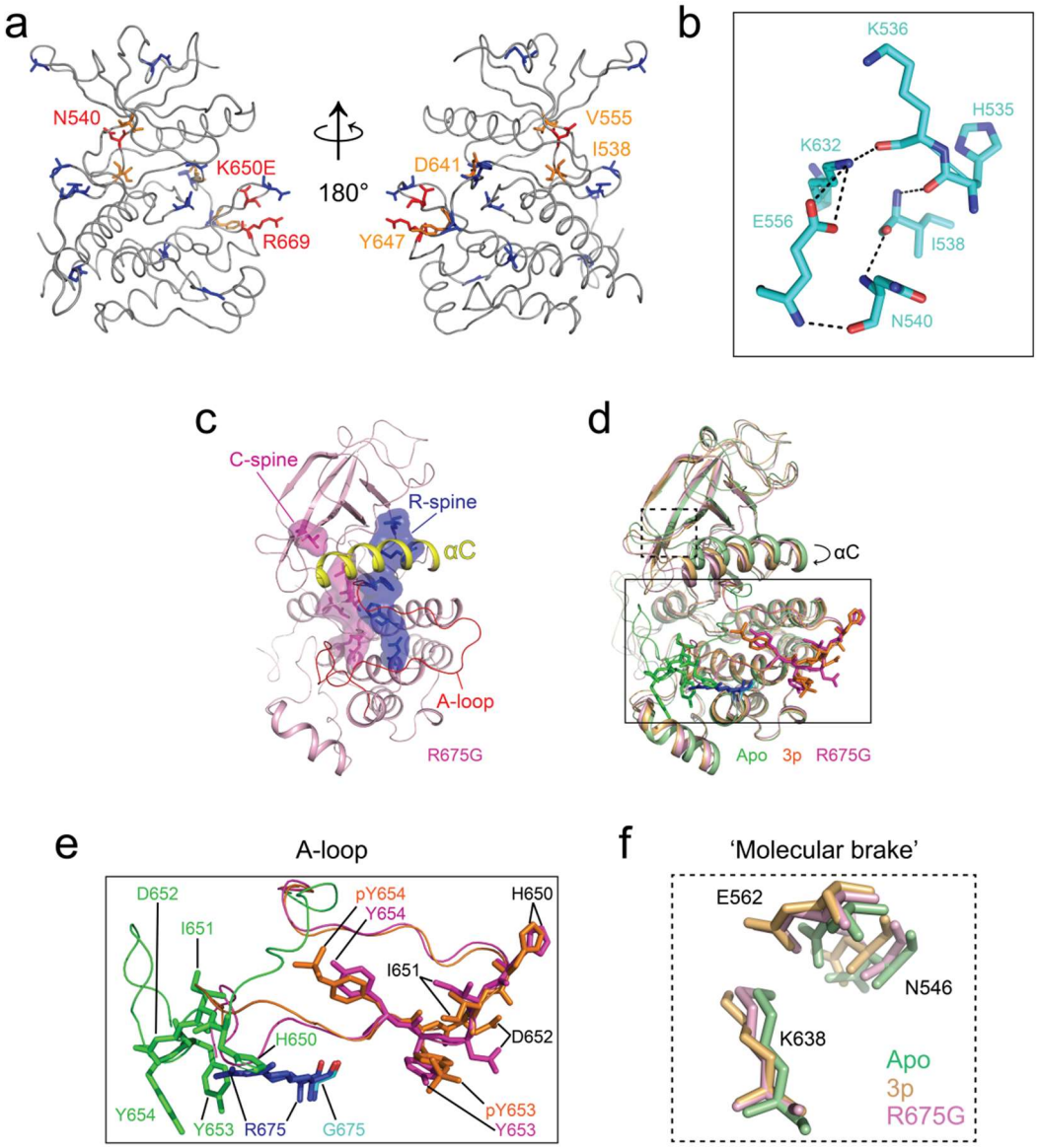
process and to date, no mutations have been identified that can fully activate the kinase as much as fully activated FGFR-WT.

3.3.7 Mutations differently affect KD phosphorylation kinetics and ATP affinity

To further investigate the impacts of these diverse KD mutations on kinase activity, the mutant KDs were next tested in two different experiments using the *in vitro* kinase activity assay to assess the impact of mutations on phosphorylation rates, kinetics and substrate binding.

First, a subset of FGFR3 variant KDs was assayed for substrate phosphorylation over a 60 minute time course. A single concentration of each of the nine FGFR3 mutants, ATP and polyE_4Y_1 substrates was used respectively, similar to the experimental design of the phosphorylation screens. Kinase reactions were stopped at different time points between 0-60 minutes (see section 2.9 for full experiment details) and compared to FGFR3-2C. Data were fitted with an exponential model for pseudo-first order association kinetics (equation **3**) to calculate the KD phosphorylation rates, rate constants and other parameters and these values are listed in Table 3.3. The raw time course curves and the rate constants for all mutants are displayed in Figure 3.14a-b. The data suggests that both mutations at N540 (N540K and N540S) as well as K650E result in the largest increases in the initial rate of phosphorylation compared to FGFR3-2C, followed by D641G/N and K650N. The R669G mutation only has a slightly elevated initial rate of reaction compared to FGFR3-2C but results in the largest total phosphorylation response over the course of the 60 minute assay, closely followed by K650E and N540K. The other mutation at this position (R669Q) has a much reduced initial rate of reaction, but both mutations at R669 are predicted by the kinetic model to cause the largest ultimate maximum phosphorylation response by 60 minutes (Table 3.3). In contrast, G697C demonstrates a reduced maximum response, has a reduced reaction rate and exhibits a similar rate constant to FGFR3-2C. These data are therefore consistent with the results from phosphorylation experiments in sections 3.3.4 and 3.3.5 implicating mutations N540K, K650E and R669G as the most activating mutations while hotspot G697C still does not appear to be activating.

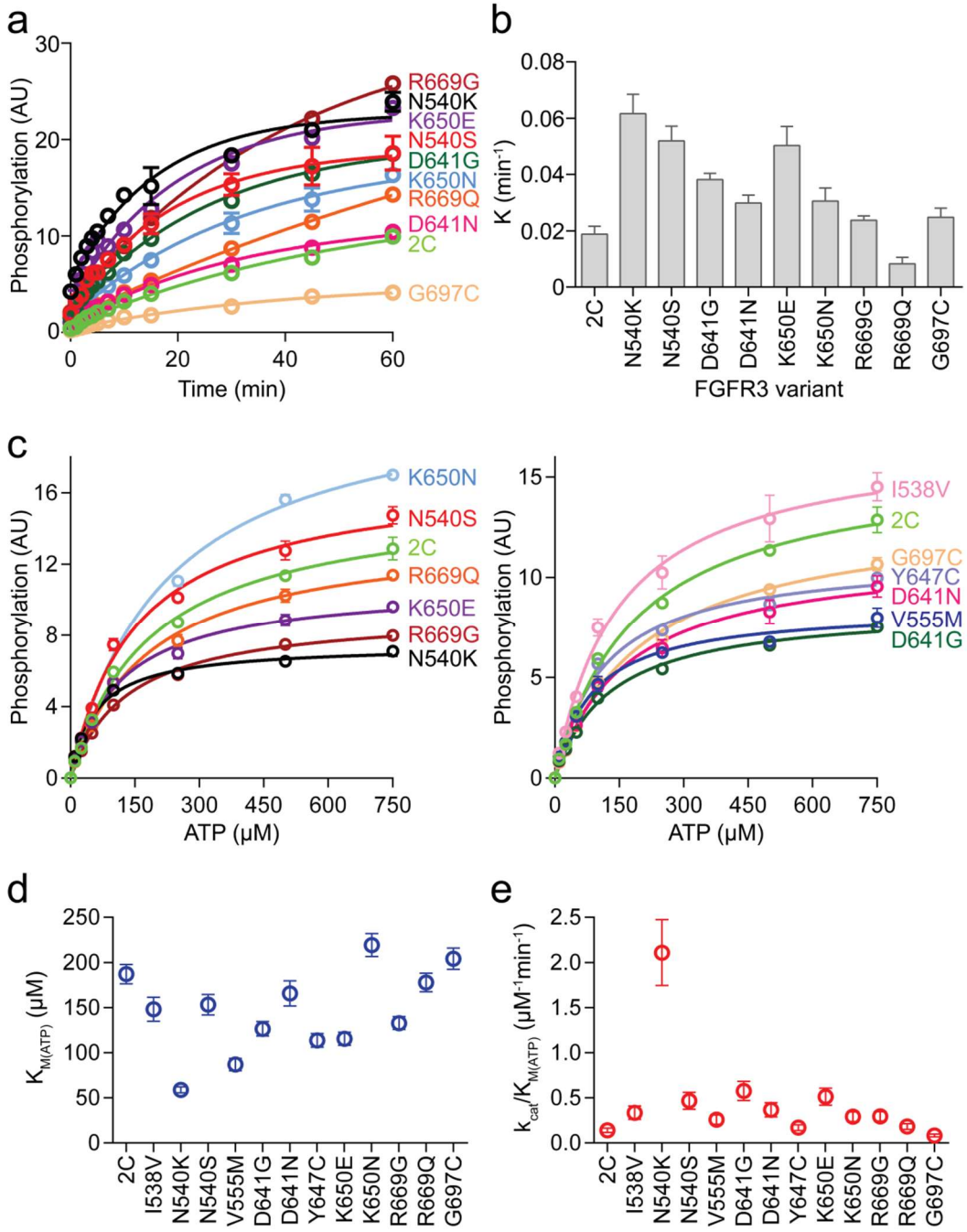
Next, the same subset of activating mutations (as well as moderately activating I538V and V555M and non-activating hotspot G697C) was tested to asses for any differences in the ability of mutant KDs to bind to ATP. The eleven mutant KDs were assayed in ATP dose response experiments and compared to FGFR3-2C. Kinase concentrations were individually determined for each mutant from fits for the substrate phosphorylation



assays performed in section 3.3.5 to ensure they fell within the mutant-specific linear concentration range (see section 2.9 for full experimental details). Data were then fitted with the Michaelis-Menten equation and dose response curves are shown in Figure 3.14c, presented in two groups for visual clarity (left and right panels). From these curves, the affinity constants for ATP binding ($K_{M(ATP)}$), turnover rates and catalytic efficiencies were calculated for each variant and these are presented in Table 3.4 and Figure 3.14d-e. These data suggest that mutations impact the FGFR3 KD's affinity for ATP and turnover to differing degrees. The two mutations that caused the largest increases in ATP binding affinity were N540K and V555M, a 'molecular brake' and gatekeeper mutation respectively. These mutations occur in the kinase hinge region, where both residues normally contribute to the ATP binding pocket (Figure 3.5b).

Figure 3.13 Possible mechanisms of partial ligand independent activation employed by activating mutations. (a) Activating mutations from the kinase activity assays in this work mapped on a cartoon representation of the only available FGFR3 KD crystal structure (FGFR3-K650E) [PDB: 4K33] (Huang et al., 2013b) (coloured grey). Residues at which mutations were not found to be activating are shown as blue sticks in both panels, rotated relative to each other by 180° around the vertical axis as indicated. Positions at which highly activating mutations occur are highlighted in red and labelled (left panel) and remaining positions at which mutations were moderately activating are highlighted in orange and labelled (right panel). **(b)** Close-up view of the kinase hinge region of the same FGFR3 structure as in **a** but in a different orientation, showing the six residues of the extended 'molecular brake' (shown as sticks and coloured cyan and by atomic element with nitrogen atoms in navy blue and oxygen in red), including N540 and I538. Activating mutations at these positions directly disengage the 'molecular brake' by disrupting hydrogen bonds (shown as dashed lines and coloured black). **(c)** Cartoon representation of the crystal structure of active FGFR1-R675G (corresponding to FGFR3-R669G) [PDB: 5FLF] (Patani et al., 2016) (coloured pink), showing the αC-helix (yellow) and the A-loop (red) in the active conformation. Regulatory and catalytic R- and C- spines are also ordered in this structure (spine residues shown as sticks against a surface and coloured navy blue and bright pink for R- and C-spines respectively). **(d)** Overlay of the FGFR1-R675G structure in **c** with inactive (apo) FGFR1 [PDB: 4UWY] (Patani et al., 2016) and active phosphorylated FGFR1 (3p) [PDB: 3GQI] (Bae et al., 2009) coloured pale green and pale orange respectively. The difference in positioning of the αC-helix between the apo structure and the 3p and R675G structures is indicated. **(e)** Close-up view of the large inset in **d** showing the position of the A-loop in the same three structures as in **d**, namely apo (green), 3p (orange) and R675G (bright pink). Key residues are shown as sticks, demonstrating that the positions of residues in the R675G structure match more closely to the same residues in the 3p than the apo structure. **(f)** Close-up view of the region corresponding to the small inset in **d** (but in a different orientation) showing the 'molecular brake' triad of residues in the same three structures as in **d** and in the same colours. The positions of the three residues in the R675G structure are at positions intermediate between the same residues in the apo and 3p structures, showing a partial disengagement of the 'molecular brake'.

Therefore, it is hypothesised that specific mutations at these positions could alter the stereochemistry within this pocket either in favour of ATP binding, or perhaps even in favour of ADP release. This might also explain why N540S, a replacement at the same position as N540K, differently impacts ATP binding to a lesser extent than N540K. Interestingly, the corresponding mutation in FGFR2 (N549K) has also been shown to cause a reduction in the binding affinity for ATP-competitive inhibitors (Byron et al., 2013), while gatekeeper mutations FGFR3-V555M and the corresponding mutation FGFR1-V561M also affect the binding of ATP-competitive inhibitors to differing degrees (Bunney et al., 2015), in both cases leading to intrinsic resistance to these ligands. These effects, as well as the increased and mutation specific ATP binding affinities observed in this work, are likely a consequence of the same phenomenon whereby replacement of the sidechain of these key residues in the ATP binding pocket improves ATP binding while concurrently diminishing the binding of ATP-competitive inhibitors in this site due to mutation-induced stereochemical effects. The N540K mutation is additionally unique in that, while most mutations only modestly affect the overall catalytic efficiency of the FGFR3 KD, N540K has a large impact by greatly increasing catalytic efficiency compared to FGFR3-2C (Figure 3.14e). This is likely the result of the combination of its increased initial reaction rate, turnover and increased ATP binding affinity, the latter of



which could improve phosphoryl transfer through improved ATP binding in the active site. It is unclear from these experiments how this mutation might affect release of ADP from the ATP binding site following phosphoryl transfer, a catalytic step thought to be rate limiting for kinase *trans*-autophosphorylation (Adams, 2001). Further experiments to assess mutation specific ADP binding will be required to fully understand this process for the FGFR KD. Nonetheless, the mechanism of activation by N540K appears to be under kinetic control, unique among the panel of mutants tested in this work.

Figure 3.14 Determination of rates of phosphorylation, affinity constants for ATP and catalytic efficiencies of activating mutations. (a) Time course experiments performed using 0.5 μM FGFR3-2C, 28.6 μM polyE_4Y_1, 500 μM ATP and time points between 0-60 minutes. Each data point was measured in triplicate and error bars represent the resulting standard deviation at each time point. Data are representative of four independent experiments performed using FGFR3-2C, and of two independent experiments performed for some of the variants. Data for G697C is coloured pale orange, and colouring for other variants is the same as in Figures 3.10 and 3.11. Data were fitted to a one phase exponential model for pseudo-first order association kinetics (equation **3**) and used to calculate the phosphorylation rate constants (K) for the reactions that are displayed as a histogram in **b**. AU – arbitrary units. **(b)** Rate constants calculated from reactions in **a**. Error bars represent standard errors calculated from the non-linear fits in **a**. Calculated values for rate constants and other parameters from the fits in **a** are also presented in Table 3.3. **(c)** ATP titration experiments performed using variable concentrations of FGFR3-2C (calculated from the data in Figure 3.11a-b, see section 2.9 for full experimental details), 42.9 μM polyE_4Y_1, 0-750 μM ATP and 10 minute reactions for variants at 2C and positions N540, K650E and R669G (left panel) and variants at 2C and positions I538, V555, D641, Y647 and G697 (right panel). Colouring for each variant is the same as in **a**. Each data point was measured in triplicate and error bars represent the resulting standard deviation for each concentration of ATP. Data were fitted to the Michaelis-Menten equation (equation **4**) and this was used to calculate affinity constants for ATP ($K_{M(ATP)}$) and catalytic efficiencies ($k_{cat}/K_{M(ATP)}$) for the reactions that are displayed as scatter plots in **d** and **e**. Scatter plots of **(d)** $K_{M(ATP)}$ and **(e)** $k_{cat}/K_{M(ATP)}$ calculated from reactions in **c**. Error bars represent standard errors calculated using fitted data in **c**. Calculated values for affinity constants, catalytic efficiency and other parameters from the fits in **c** are also presented in Table 3.4. For full experimental details see section 2.9.

Table 3.3 Maximum response [AU, arbitrary units], initial reaction rate (calculated at 3 minutes), rate constant (K), time constant (1/K) and 'half-time' $T_{1/2}$ (ln2/K) of variant FGFR3 KDs (± standard error where applicable) measured using *in vitro* kinase activity assays, to complement data in Figure 3.14a-b, and calculated using equation **3**.

FGFR3 variant	Maximum response [AU]	Initial reaction rate [AU/min]	K [min^{-1}]	1/K [min]	ln2/K [min]
2C	13.18 ± 1.14	0.377 ± 0.013	0.0189 ± 0.0028	53.0	36.7
N540K	17.27 ± 0.63	1.565 ± 0.129	0.0616 ± 0.0069	16.2	11.3
N540S	16.55 ± 0.56	1.191 ± 0.161	0.0520 ± 0.0053	19.3	13.3
D641G	17.89 ± 0.39	0.868 ± 0.015	0.0383 ± 0.0022	26.1	18.1
D641N	10.98 ± 0.45	0.558 ± 0.034	0.0299 ± 0.0028	33.4	23.2
K650E	19.75 ± 0.90	2.071 ± 0.078	0.0503 ± 0.0068	19.9	13.8
K650N	16.83 ± 1.18	1.040 ± 0.078	0.0305 ± 0.0047	32.8	22.7
R669G	30.11 ± 1.03	0.954 ± 0.071	0.0238 ± 0.0016	42.1	29.2
R669Q	32.59 ± 7.08	0.022 ± 0.016	0.0083 ± 0.0023	120.6	83.6
G697C	4.89 ± 0.31	0.138 ± 0.035	0.0249 ± 0.0032	40.2	27.9

Table 3.4 ATP affinity constants ($K_{M(ATP)}$), turnover rates (k_{cat}) and catalytic efficiencies (k_{cat}/K_M) of variant FGFR3 KDs (± standard error where applicable) measured using *in vitro* kinase activity assays, to complement data in Figure 3.14c-e, and calculated using equation **4**. Concentrations of FGFR3 variants to be used in these assays (column 2) were calculated from each saturation binding curve using data in Figure 3.11a-b and equation **2**. For full experimental details see section 2.9. It should be noted that the KD concentrations in column 2 have uniquely been assigned a theoretical error, as it is not possible to measure nanomolar protein concentrations using standard instrumentation.

FGFR3 variant	Concentration used in assays [nM]	$K_{M(ATP)}$ [μM]	k_{cat} [min^{-1}]	k_{cat}/K_M [μM min^{-1}]
2C	600 ± 60.0	187.1 ± 10.7	26.4 ± 0.5	0.141 ± 0.002
I538V	343 ± 34.3	148 ± 13.2	49.7 ± 1.5	0.335 ± 0.008
N540K	60 ± 6.0	58.8 ± 3.4	124.1 ± 1.9	2.109 ± 0.039
N540S	238 ± 23.8	153.2 ± 11.3	71.7 ± 1.9	0.468 ± 0.003
V555M	377 ± 37.7	86.9 ± 6.7	22.5 ± 0.5	0.259 ± 0.002
D641G	177 ± 17.7	126.4 ± 8.1	72.7 ± 1.5	0.575 ± 0.005
D641N	186 ± 18.6	165.7 ± 14.0	60.8 ± 1.8	0.367 ± 0.007
Y647C	575 ± 57.5	113.7 ± 7.5	19.3 ± 0.4	0.170 ± 0.001
K650E	183 ± 18.3	115.5 ± 7.2	59.3 ± 1.1	0.514 ± 0.006
K650N	345 ± 34.5	219.3 ± 12.8	63.8 ± 1.5	0.291 ± 0.003
R669G	238 ± 23.8	132.8 ± 7.1	39.2 ± 0.7	0.295 ± 0.006
R669Q	430 ± 43.0	177.8 ± 10.2	32.3 ± 0.7	0.182 ± 0.003
G697C	795 ± 79.5	204.2 ± 11.7	16.8 ± 0.4	0.082 ± 0.001

3.3.8 Kinase activation is not linked to mutation-induced KD instability

It was hypothesised that variants might impart differing degrees of local and/or global structural modifications to the FGFR3 KD which might in turn result in increased conformational plasticity of the inactive KDs. This might 'prime' individual variants for kinase activation to differing extents by shifting the conformational equilibrium from inactive to more active states. Therefore, to determine experimentally whether there might be a link between thermal stability and activation status of the KD variants, thermal melting experiments were performed on the panel of purified variant FGFR3 KD proteins using a thermal shift assay with fluorimetric readout. Protein samples in triplicate were mixed with a fluorescent hydrophobic dye (SYPRO® Orange) and subjected to a temperature gradient from 10-95°C increasing at 0.5°C intervals. Curve fitting was performed with an appropriate custom equation (for full experimental details see section 2.20.1)) and the melting temperatures were extracted from the resulting melting curves (Figure 3.15a-b). All curves showed a clear melting transition allowing determination of

the melting temperatures calculated at the point of inflection. The melting temperature for FGFR3-2C (curve highlighted in green) was determined to be 43.8°C. Most FGFR3 variants appeared to have similar melting temperatures to FGFR3-2C of 40°C or higher. No mutations appeared to cause a large increase in stability. The highly activating mutations K650E (located in the A-loop) and R669G and moderately activating R669Q (latter two located in the αEF-β12 loop) had the largest standard errors in their calculated melting temperatures, suggesting an intrinsic conformational heterogeneity in these activating mutant KDs that presents as a wider 'range' of melting temperatures.

Variants causing the most striking reductions in melting temperatures of the FGFR3 KD to below 37°C (marked by the dashed line on Figure 3.15b) were E466K, I538F and N540K with measured melting temperatures of 34.4°C, 32.9°C and 34.9°C. A different replacement to V at position I538 also caused a moderate reduction in the melting temperature to 39.1°C, however the replacement to S at position N540 does not appear to alter melting temperature. Residue E466 is located in the juxtamembrane region while I538 and N540 are located in the αC-β4 loop; all in the N-lobe of the split KD (Figure 3.15c). N540K, I538F and I538V are respectively part of or adjacent to the 'molecular brake' triad of residues that stabilise the inactive KD in an autoinhibitory state through a hydrogen bonding network. The replacement to K at position N540 directly disrupts these hydrogen bonds, releasing autoinhibition and destabilising the inactive state as shown for corresponding mutations in FGFR2 (Chen et al., 2007b). Mutations at residue I538 adjacent to the 'molecular brake' and part of the 'DFG-latch' contribute to destabilisation by similar disruptions of 'molecular brake' hydrogen bonds, including possible steric effects that appear to increase in impact with the size of the replacement side chain such as for I538F, possibly affecting the local structure of the 'DFG-latch'. It is not clear how mutation at residue E466 might exert its destabilising effect. It is thought that the FGFR juxtamembrane region might mediate dimerisation through a mechanism similar to EGFR (see section 1.5.3) by juxtapositioning the KDs to form *trans*-phosphorylation competent asymmetric dimers (Chen et al., 2008, Peng et al., 2009, Bae et al., 2010, Bocharov et al., 2013). It is not known whether E466 is involved in FGFR dimerisation, but this residue is conserved across all four FGFRs and introduction of a similar mutation to the corresponding residue in FGFR1 has also been shown to reduce the thermal stability of the FGFR1 KD (unpublished data). In addition, mixing the variant FGFR3 KDs with FGFR inhibitors dramatically increases the melting temperature (and therefore stability) of all three destabilising variants (unpublished data). The molecular mechanisms behind these effects remain to be elucidated.

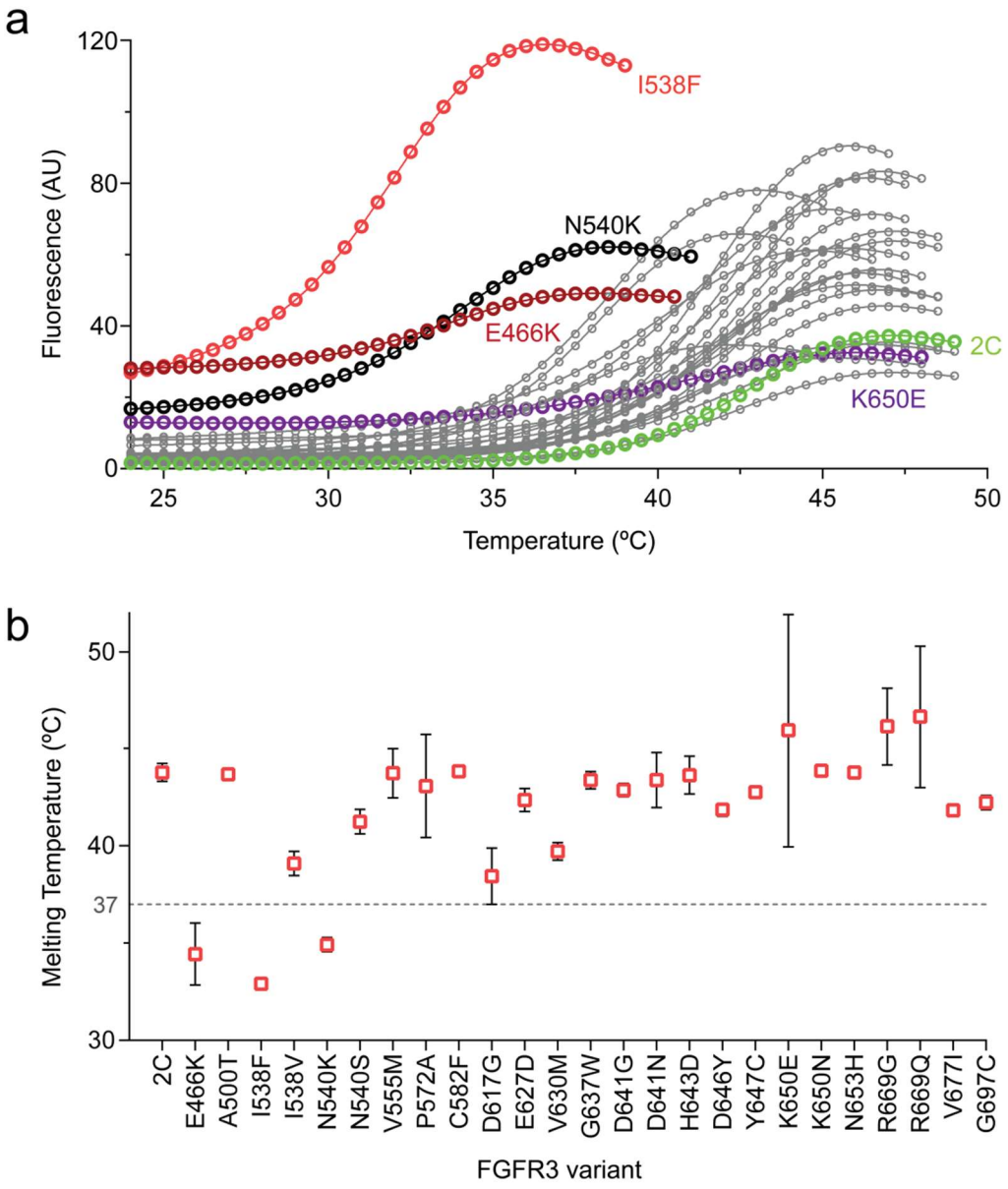

a
120
80
40
0
Fluorescence (AU)
I538F
N540K
E466K
2C
K650E
25
30
35
40
45
50
Temperature (ºC)
b
50
40
37
30
Melting Temperature (ºC)
2C
E466K
A500T
I538F
I538V
N540K
N540S
V555M
P572A
C582F
D617G
E627D
V630M
G637W
D641G
D641N
H643D
D646Y
Y647C
K650E
K650N
N653H
R669G
R669Q
V677I
G697C
FGFR3 variant
c
E466
N540
I538
K650E

Figure 3.15 Thermofluor assay measuring thermal stability of FGFR3 KDs. (a) Thermal shift experiments performed using 1 mg/mL FGFR3 variants mixed 1:500 with Sypro® Orange fluorescent dye (Thermo Fisher Scientific). A temperature gradient was employed between 10-95°C increasing at 0.5°C intervals (10 seconds each) for each variant in triplicate reactions. Data were analysed using a custom Boltzmann sigmoid equation with linear baselines. Three mutations whose melting curves deviated most strongly from 2C are highlighted in different colours and labelled (E466K, I538F and N540K). The melting curve for activating mutation K650E is also highlighted for comparison. Highlighted curves are coloured as I538F (light red), E466K (maroon) and other variants as in Figures 3.10, 3.11 and 3.14. All error bars are omitted for clarity. **(b)** Melting temperatures for FGFR3 KD variants calculated from data fitted in **a**. Error bars represent standard error calculated using fitted data in **a**. The dashed line at 37°C marks the mean human physiological temperature. **(c)** Positions of all mutations assayed (shown as sticks and coloured blue) marked on a cartoon representation of the only available FGFR3 KD crystal structure (FGFR3-K650E) [PDB: 4K33] (Huang et al., 2013b) (coloured grey). The positions at which the three most destabilising mutations from **a** and **b** occur are mapped on this structure and highlighted, as is the activating mutation K650E for comparison (shown as sticks against a surface and coloured as in **a**). For full experimental details see section 2.20.1.

Overall, of these destabilising variants, N540K was the only highly activating mutation, I538F and I538V were moderately activating and E466K was not an activating mutation in the *in vitro* kinase activity assays. Therefore there appears to be no link between the extent of kinase destabilisation by FGFR3 mutations and activation status. However, destabilisation of the inactive KD might play a mechanistic role in activation of some mutants such as N540K and mutations at position I538, both of which contribute to the 'molecular brake' network of autoinhibitory hydrogen bonds.

3.3.9 Do activating mutations affect cellular transformation and signalling?

The kinase activity assays provided a ranked list of variants that could act as potential oncogenic drivers. However, it is possible that some mutations do not directly impact activity of the KD under the conditions of the *in vitro* assay design. For example, one mutation which has previously been suggested to be highly activating in a cellular setting and in the context of the full length receptor is the hotspot mutation G697C (Zhang et al., 2005) which showed no gain-of-function in any kinase assays using purified and soluble KDs in this work. Conversely, it could be that many more mutations were activating in the kinase activity assays due to the increased degrees of freedom available to KDs in solution to *trans*-autophosphorylate when not restricted in space by the transmembrane and juxtamembrane domains as they would be in cells in the context of full length receptors. Therefore, the panel of FGFR KD mutations was next analysed in complementary experiments in the context of full length FGFR3 in a cellular setting to determine whether the patterns of kinase activation observed *in vitro* could also be detected in cells.

3.3.10 Generation of a panel of FGFR3 variant stable cell lines

A retroviral vector comprising full length FGFR3-WT was generously provided by the Knowles group (Tomlinson et al., 2005). The alternatively spliced FGFR3 b splice isoform was selected for these experiments (rather than the canonical c splice isoform) because FGFR3 mutations occur most frequently in urothelial and other epithelial tissues in which the FGFR b splice isoform is more commonly expressed (Figure 3.1 and section 1.4). However, to minimise discrepancies of residue numbering between isoforms (which differ by two amino acids between b and c splice isoforms), residue numbering of mutations in this full length FGFR3b construct follow the canonical numbering of the c splice isoform (as in this entire work so far). Point mutations were introduced into this pFB-FGFR3b backbone using SDM and the constructs were sequenced to confirm successful incorporation of mutations (for full methods, see sections 2.1.5 and 2.2.1). This plasmid library of FGFR3-WT and mutant receptors was used to generate stable cell lines in NIH 3T3 cells (for full experimental details, see section 2.11). Two independent cell lines were made for each FGFR3 variant and WT. Clonal cell lines were not made as it was thought this might not accurately represent the genetic heterogeneity seen in tumours, for example in urothelial carcinoma (Tomlinson et al., 2007a). Additionally, variants I538F, I538V and G637W were not included in cellular analyses because these mutations were selected for inclusion in the panel of mutations at a later date than when these experiments were performed.

3.3.11 Morphological assessment of transformation by FGFR3 variants

NIH 3T3 cell lines stably expressing each of the panel of FGFR3 constructs were first screened for morphological signs of cellular transformation (Figure 3.16). Stable cell lines were seeded at low density and imaged at subconfluence after approximately 48 hours. An NIH 3T3 cell line stably expressing the known oncogene Ras V12, also known to transform NIH 3T3 cells, was generously provided by the Knowles group and used as a positive control for transformation. Transformed fibroblasts display a characteristic loss of contact-inhibition, their cell morphology becomes more ‘spindle-like’ and the cell edges appear ‘raised’ and therefore shiny under the phase contrast microscope compared to control fibroblasts which appear ‘flat’ and have a dull appearance (Figure 3.16a). Under the conditions of this screen, the only mutations clearly exhibiting these markers of transformation were N540K and K650E. In particular, cells containing FGFR3-G697C did not appear morphologically transformed in this cell line, and neither did any cell line containing moderately activating mutations from the *in vitro* kinase activity assays. Surprisingly, the cell line containing FGFR3-R669G, the mutation which was most activating in the kinase activity assays, did not appear to transform NIH 3T3 cells.

Interestingly however, all three of these most activating mutations from the kinase activity assay displayed increased FGFR3 expression, as did G697C (Figure 3.16b). From this initial experiment it was not known whether this was due to heterogeneity of the stable cell lines due to different efficiencies of gene integration into the genome during generation of the stable cell lines, or whether it might be the result of a feed-forward mechanism in response to mutation-induced signal transduction.

In order to eliminate the possibility that the transforming effects of activating mutations were due to increased receptor expression, a second FGFR3-WT stable cell line (a gift from Sarah Williams) displaying elevated FGFR3 expression similar to the N540K and K650E cell lines (Figure 3.17a, left panel) was compared with these two mutant cell lines. Increased WT receptor expression in these cells did not morphologically transform them; the cells appeared morphologically similar to the WT cell line with lower receptor expression (Figure 3.17a, middle panel, and 3.17b). Additionally, two independent K650E stable cell lines (the second also a gift from Sarah Williams) with different receptor expression levels (Figure 3.17a, right panel) both transformed NIH 3T3 cells (Figure 3.17b), suggesting that transformation is not a result of increased receptor expression. In any case, this type of morphology screen is subjective and qualitative, so further quantitative experiments were designed to elucidate whether the elevated receptor expression levels were mutation-induced or a transfection artefact, and to further characterise the mutations in cells with respect to their effects on cellular transformation and signal transduction.

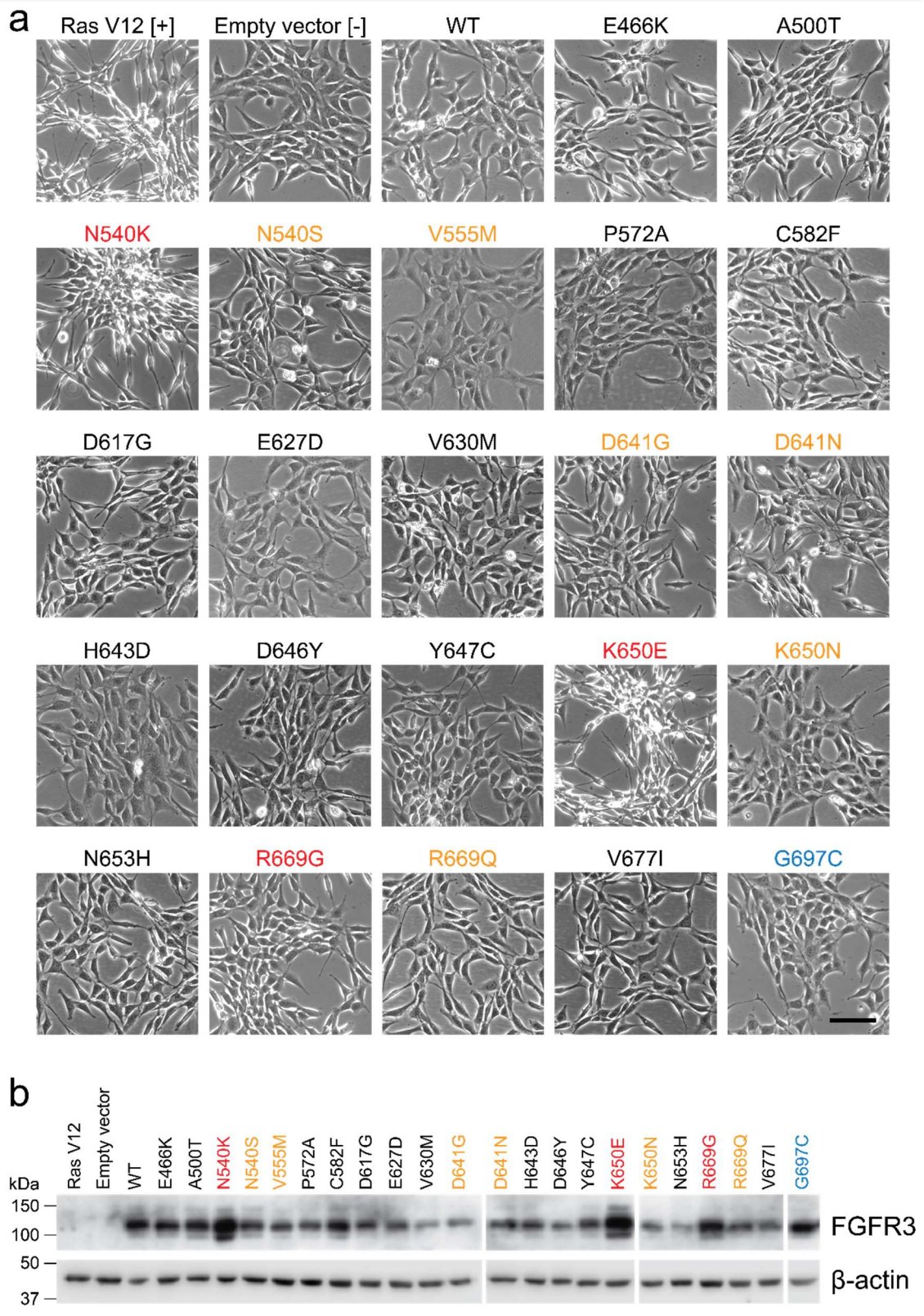

a
Ras V12 [+]
Empty vector [-]
WT
E466K
A500T
N540K
N540S
V555M
P572A
C582F
D617G
E627D
V630M
D641G
D641N
H643D
D646Y
Y647C
K650E
K650N
N653H
R669G
R669Q
V677I
G697C
b
kDa
150
100
50
37
Ras V12
Empty vector
WT
E466K
A500T
N540K
N540S
V555M
P572A
C582F
D617G
E627D
V630M
D641G
D641N
H643D
D646Y
Y647C
K650E
K650N
N653H
R669G
R669Q
V677I
G697C
FGFR3
β-actin

Figure 3.16 Assessment of mutation-induced morphological transformation of NIH 3T3 cells. (a) Images of stable NIH 3T3 cell lines generated with full length FGFR3b-WT or variants. Mutations that were highly (red) or moderately (orange) activating, and the non-activating hotspot mutation G697C (blue) in the *in vitro* kinase activity assays are highlighted. A stable cell line expressing the transforming Ras V12 oncoprotein (a gift from the Knowles group, University of Leeds) is also shown (positive control, [+]). A stable cell line generated using the empty vector backbone used for generation of all FGFR3 constructs is also shown (negative control, [-]). Stable cell lines were seeded at low density and imaged at subconfluence approximately 48 hours later. The scale bar represents 100 μM. **(b)** Relative expression of FGFR3 variants in each cell line in **a**, shown using immunoblots performed with the indicated antibodies. The gaps between panels represent the same blot with all panels imaged together, but with samples run on different SDS-PAGE gels or excluding intervening lanes with mutants no longer included in this work. For full experimental details see sections 2.4, 2.5 and 2.10.

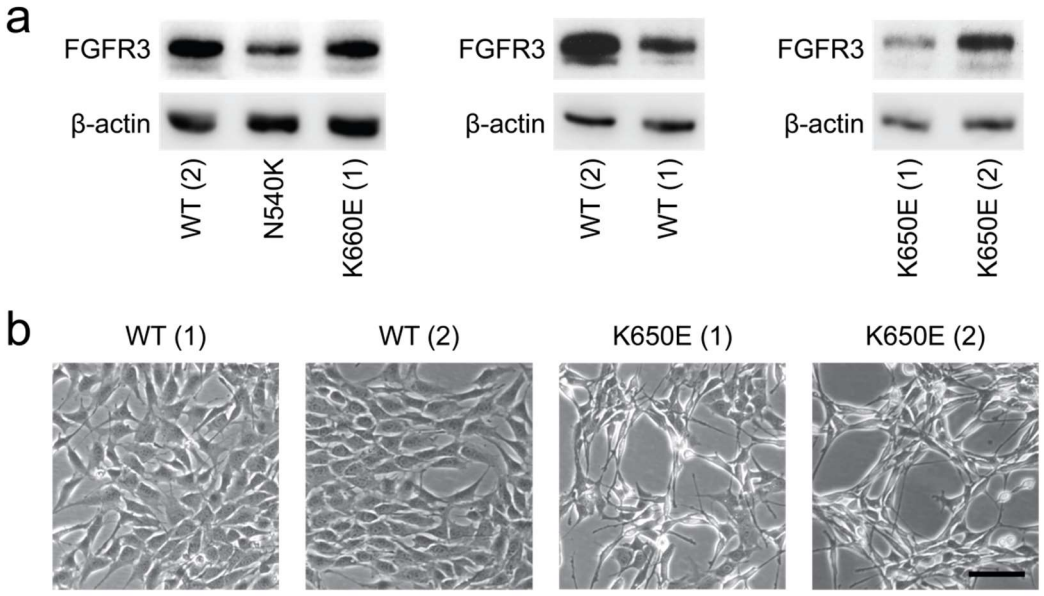


Figure 3.17 Transformation of NIH 3T3 cell lines occurs independently of the level of FGFR3 expression. (a) Expression of FGFR3 variants in indicated stable cell lines, shown using immunoblots performed with the indicated antibodies. Immunoblots in the left, middle and right panels are from independent experiments. Stable cell lines WT (2) and K650E (2) were a gift from Sarah Williams. **(b)** Images of stable NIH 3T3 cell lines used for immunoblots in **a**. Stable cell lines were seeded at low density and imaged at subconfluence approximately 48 hours later. The scale bar represents 100 μM. For full experimental details see sections 2.4, 2.5 and 2.10.

3.3.12 Effects of mutations on anchorage independent growth and cell signalling

A subset of FGFR3 variants was selected for analysis in cellular assays using the stable NIH 3T3 cell lines. The three most highly activating mutations from the kinase activity assays (N540K, K650E and R669G) were selected alongside the hotspot G697C mutant that appeared to be non-activating in the *in vitro* assays. These four mutant cell lines were assayed in parallel and compared with FGFR3-WT. Three types of experiments were performed: anchorage independent growth in soft agar, immunoprecipitation (IP) of FGFR3 with subsequent immunoblotting and immunoblot experiments performed directly on cell lysates. For experiments requiring cell lysates, cells were grown to subconfluence and then were either starved overnight in serum-free medium or harvested directly without starvation, and cell lysis was performed *in situ*. Lysates were then either used immediately for further experiments (following a total protein quantification step) or frozen until next use. Detailed methods can be found in sections 2.4, 2.5, 2.10, 2.13, 2.14 and 2.15.

The first experiment tested the ability of mutations to induce anchorage independent growth of the NIH 3T3 stable cell lines in soft agar to assess whether the mutations were transforming. Each stable cell line was seeded in triplicate in freshly prepared soft agar, cells were fed weekly and colonies were counted and imaged after two weeks (for full experimental details see section 2.13). Cells used for these experiments were also imaged for their morphological transformation status during anchorage dependent growth (Figure 3.18a). The soft agar assay showed that highly activating and morphologically transforming mutations N540K and K650E were also capable of inducing anchorage independent growth (Figure 3.18b-c). This has been previously shown for the positive control K650E (Chesi et al., 2001, Ronchetti et al., 2001, Agazie et al., 2003, di Martino et al., 2009) but was shown for the first time for N540K or any mutation at this position in FGFR1-4, although it has recently been shown for the corresponding mutation N549K in FGFR2 (Kwak et al., 2015). However, the highly activating mutation R669G and the hotspot mutation G697C appeared to be incapable of inducing anchorage independent growth in soft agar (Figure 3.18b-c), in agreement with their lack of morphological transformation of NIH 3T3 cells (Figure 3.16a).

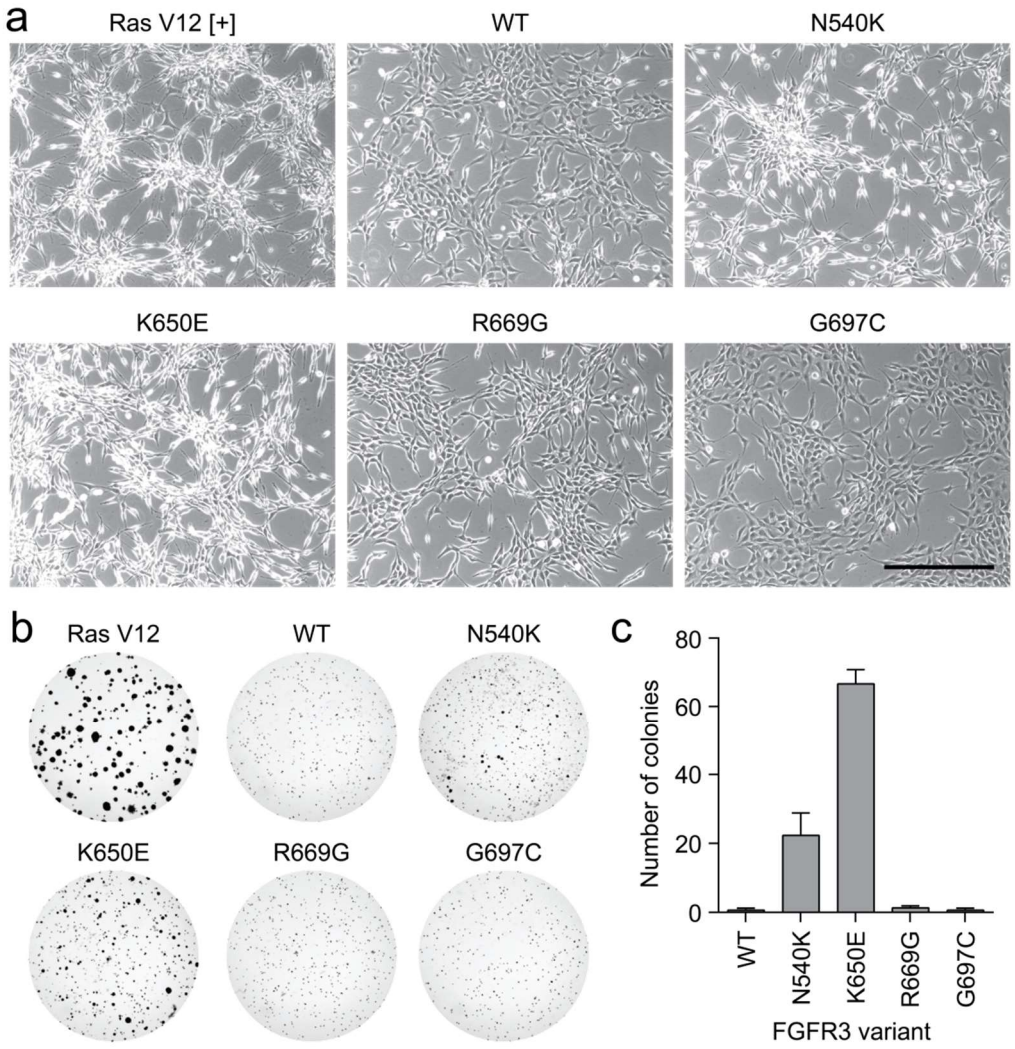


Figure 3.18 Anchorage independent growth assay. (a) Images of a subset of the same stable NIH 3T3 cell lines as in Figure 3.16 prior to their use in anchorage independent growth experiments. Stable cell lines were seeded at low density and imaged at subconfluence after approximately 48 hours. The scale bar represents 200 μm. These cells were then seeded at 5×10^3 cells per well of a 6-well plate in medium containing 0.4% w/v agarose on a base of medium containing 0.8% w/v agarose. Cells were fed weekly with medium containing 0.4% w/v agarose. After two weeks, cells were stained for 24 hours with 0.3% w/v p-iodonitrotetrazolium violet and imaged. **(b)** Representative views of binary masks applied to images of colonies growing in the soft agar (black against a white background) during automated colony counting, applied using ImageJ. **(c)** Number of colonies counted for each FGFR3 variant from triplicate wells in the soft agar assay. Colonies were counted electronically using ImageJ and were then manually checked. Error bars represent the standard deviation for each FGFR3 variant. Colonies counted for Ras V12 (positive control) are not shown on this bar chart but can be seen in **b**. For full experimental details see section 2.13.

Next, IP and immunoblotting experiments were performed to determine whether these mutations elevated the level of FGFR3 phosphorylation in the context of the full length receptors, as well as hyperactivation of downstream signalling pathways (for full experimental details see section 2.10, 2.14 and 2.15). It was found to be challenging to extract similar quantities of FGFR3-WT and each variant by IP, partly due to the variable expression levels observed in each cell line. This particularly affected the experiments including R669G, as this mutant protein could only be extracted in very small quantities (Figure 3.19a). Therefore, following IP of FGFR3-WT and variants from all lysates, R669G was compared with FGFR3-WT in separate immunoblots to the other three variants using a comparable quantity of IP-extracted FGFR3-WT (Figure 3.19b). These experiments showed that all three activating mutations (N540K, K650E and R669G), but not G697C, displayed elevated receptor autophosphorylation compared to WT in cells in the absence of ligand stimulation, in agreement with the *in vitro* kinase activity data.

Assessment of the impact of these mutations on downstream signalling using non-IP extracted lysates (see sections 2.10 and 2.14 for detailed methods) once again showed that all three activating mutations, but not G697C, activated downstream signalling pathways compared to FGFR3-WT as evidenced by stronger phosphorylation of ERK (pERK) and AKT (pAKT) compared to WT (Figure 3.19c-d). This outcome was observed for R669G despite its relatively low expression level compared to N540K and K650E. Interestingly, the two different K650E cell lines (with varied expression levels, see Figure 3.17a, right panel) appeared to differently activate either ERK signalling or AKT signalling, but did not fully activate both at the same time. Additionally, in the K650E cell line in which pERK was downregulated, the expression of PLCγ1 was also reduced compared to all other cell lines. This suggests that K650E might be capable of switching between different downstream signalling pathways, possibly to diversify and fine-tune its range of signalling outcomes in a context dependent manner. It remains to be tested whether this kind of pathway switching might also apply to other activating mutations, as in this experiment both N540K and R669G only weakly activated pAKT alongside stronger activation of pERK (Figure 3.19c-d). Whether these signalling outcomes depend on different thresholds of receptor expression within the cells also remains to be determined.

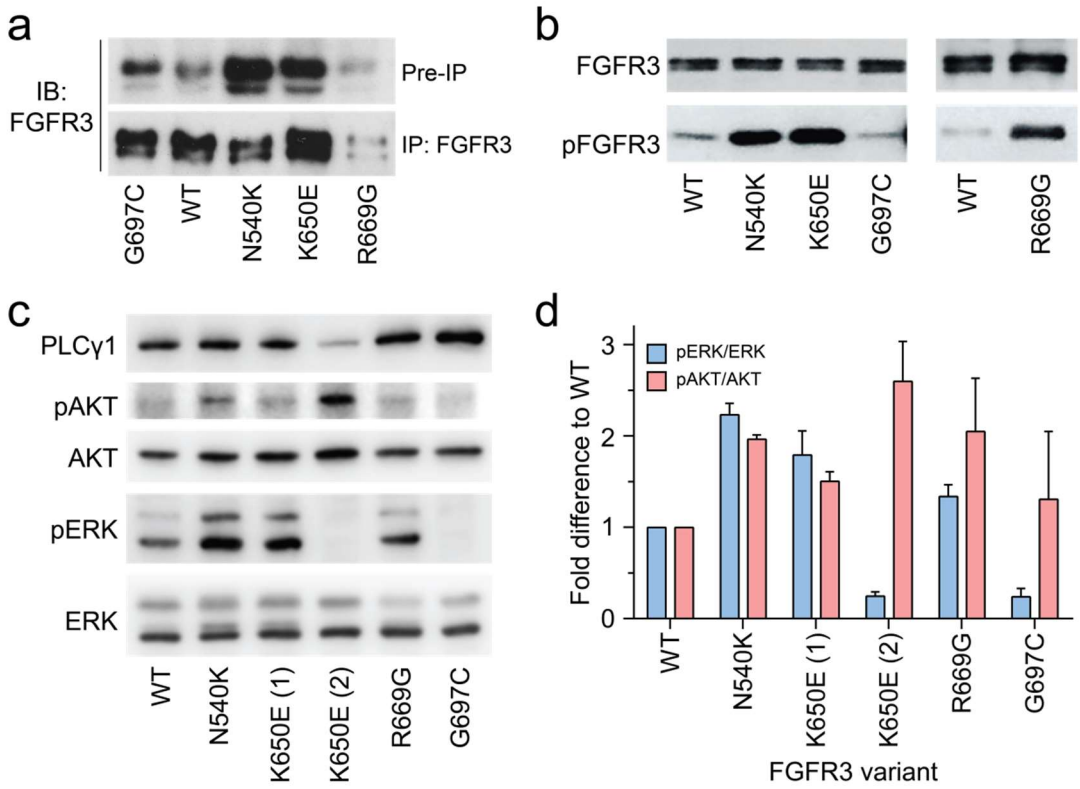


Figure 3.19 Effect of mutations on downstream signalling. (a) Immunoprecipitation (IP) of FGFR3 variant proteins from NIH 3T3 stable cell lines. Immunoblots were performed using the indicated antibody on lysates prepared from the same stable NIH 3T3 cell lines as in Figure 3.16 showing the levels of FGFR3 expression (Pre-IP, upper panel) and showing the differences in the quantities of immunoprecipitated FGFR3 variants following equal volume IPs of FGFR3 proteins from the indicated variant cell lines (IP, lower panel). **(b)** Immunoblots performed using the indicated antibodies on immunoprecipitated FGFR3 variants showing differences in the levels of phosphorylated FGFR3 variants for equal quantities of FGFR3. IP and immunoblotting for this Figure was performed by the Research and Clinical Pathology Support Unit at the UCL Cancer Institute **(c)** Analysis of ligand-independent downstream signalling pathway activation by indicated FGFR3 variants. Stable cell lines (from Figure 3.16) were seeded at low density and at subconfluence (approximately 48 hours later) they were starved overnight in serum-free medium prior to harvesting the cell lysates. Immunoblots were performed on the lysates using the indicated antibodies. Immunoblots shown are representative of two independent experiments. **(d)** Densitometric analyses of data from **c**, plotted on a bar chart as fold difference compared to WT (normalised WT=1). Error bars represent resulting standard deviation of the data from the two independent experiments. For full experimental details see sections 2.14 and 2.15.

To begin to gain further insights into possible mechanisms of activation via alteration of downstream signalling by these mutant FGFR3 receptors, an intracellular signalling antibody array experiment was performed using the same subset of FGFR3 variants. Cell lysates were added to the wells of an antibody array chip and chemiluminescent signals from each antibody 'spot' in the array (two per phospho-protein) were captured (for full experimental details see section 2.16). Raw data from the arrays is shown in Figure 3.20a, and the results of densitometric analyses performed on these data are displayed in Figure 3.20b. Since only a single experiment was performed, these are only

preliminary data; however, several interesting observations can be made from this preliminary dataset. The first is that all three activating mutations appear to activate the Ras-RAF-MEK-MAPK and the PI3K-PDK-AKT pathways in this assay, evidenced by elevated pERK and pAKT (similar to immunoblot experiments in Figure 3.19c-d) and phosphorylated proline-rich AKT substrate of 40 kDa (PRAS40) compared to WT, although G697C also appears to activate ERK and AKT in this array. Phosphorylation of BCL2-associated agonist of cell death (BAD) by N540K, K650E and G697C is also observed, suggesting anti-apoptotic signalling in these mutant cell lines. These combined data suggest that mutations increase pro-survival signalling in these cells compared to WT. In contrast, STAT3 does not appear to be activated here, despite this being previously suggested for K650E in the context of dysplasias (Hart et al., 2000). Phosphorylation of AMPKα, the key sensor of cellular energy status (Garcia and Shaw, 2017) also appears unaffected, suggesting mutations do not appear to greatly impact cellular metabolism and/or energy homeostasis. Intriguingly, a strong reduction in active p38 is seen for N540K compared to WT, but it is unclear whether this might reflect a cause or a consequence of the mutation, and is further complicated by the fact that p38 can have context-dependent oncogenic or tumour-suppressive roles (Zou and Blank, 2017). Finally, all four mutations display phosphorylation and therefore inactivation of glycogen-synthase kinase-3 beta (GSK-3β), suggesting crosstalk with the Wnt and insulin signalling pathways. In fact, several of these potentially activated pathways exhibit substantial crosstalk within cells, and rigorous experimental testing is needed to elucidate the wider impacts of these mutations and their mechanisms of activation with regard to intracellular signalling, pathway co-activation and cellular homeostasis.

It is not known why R669G is not transforming in a cellular context, despite it being clearly activating in terms of KD phosphorylation, phosphorylation of the full length receptor and even phosphorylation of downstream effectors. Replacement at this position to glycine has not been observed in FGFR3 in cancer *per se* but has been observed in FGFR2 (Table 3.1) in which it has also been shown to activate the isolated KD *in vitro* (Chen et al., 2007b) similarly to the *in vitro* assays in this work. FGFR2-R678G was originally identified in a lung adenocarcinoma sample by the TCGA project, while a different mutation at the same position in FGFR2 (R678S) was identified in an endometrial carcinoma sample (Powell et al., 2014). In FGFR3 this position was instead found mutated to glutamine (Table 3.1), also by the TCGA project, and this has been similarly observed in FGFR1 (Seshagiri et al., 2012); both occurrences to glutamine were observed in colon adenocarcinomas. All these cancers are of epithelial origin, and so it could be that receptor activation and cellular transformation by mutations at this position only become apparent in a strictly cellular context dependent manner, and possibly in an

FGFR isoform specific fashion, which might not be detectable by the assays performed in this work using fibroblast cell lines and only one FGFR isoform.

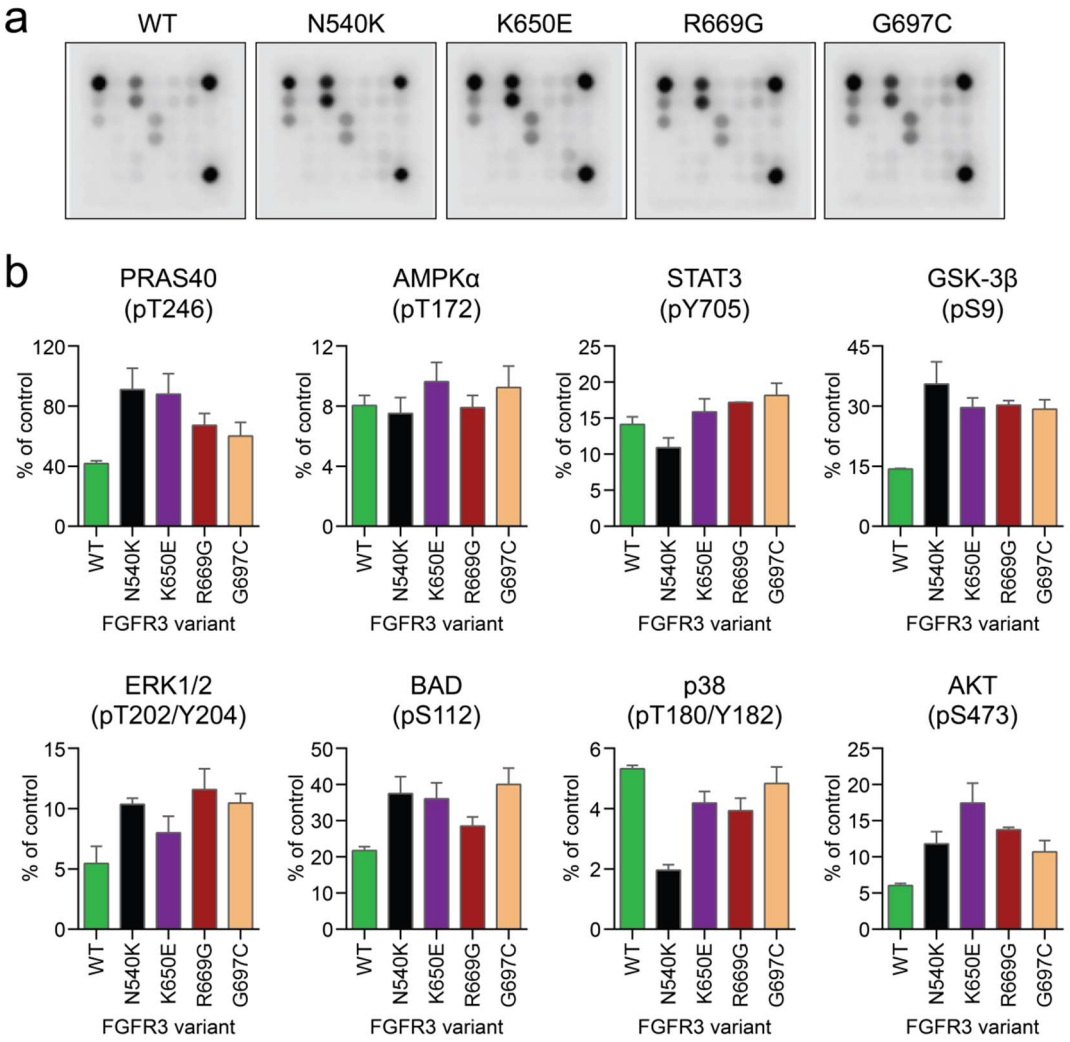


Figure 3.20 Intracellular signalling antibody array. (a) Images showing raw data from intracellular antibody arrays performed using cell lysates from the indicated FGFR3 variant stable cell lines as in Figure 3.16. Arrays were performed for each variant in parallel, in different array 'wells' on the same array chip. Two antibody 'spots' (duplicates) were present in each array 'well' for each signalling protein. Three positive control 'spots' (top left, top right and bottom right corners of each array 'well') and three negative control 'spots' (bottom left corner and the two spots directly above it in each array 'well') served as 100% and 0% signal control 'spots' respectively for each array 'well', against which all signalling data within a well were normalised. **(b)** Densitometric analyses of data from **a**, plotted on a bar chart for phosphorylated intracellular signalling proteins. Data are shown for proline-rich AKT substrate of 40 kDa (PRAS40), AMP-activated protein kinase alpha (AMPKα), signal transducer and activator of transcription 3 (STAT3), glycogen synthase kinase 3 beta (GSK-3β), mitogen-activated kinases 1 and 2 (ERK1/2), BCL-2-associated agonist of cell death (BAD), p38 mitogen-activated protein kinase (p38) and protein kinase B (AKT). Data from each array 'well' for different FGFR3 variants were normalised according to the triplicate positive and negative control 'spots' within each array 'well' as described in **a**. Error bars represent standard deviations of the data from the duplicate 'spots' of each signalling protein in each array 'well'. For full experimental details see section 2.16.

In contrast the FGFR3-G697C was not found more phosphorylated in cells compared to WT, similarly to the *in vitro* kinase assays. It did not enhance signalling via pERK and was incapable of anchorage independent growth. The combination of these results suggests this FGFR3 hotspot mutation does not appear to be relevant as an oncogenic driver, at least from the experimental conditions tested. Crucially, all the observations of this cancer mutation were detected in a single study, possibly from a single genetic population, in the FGFR3b locus in oral squamous cell carcinomas (OSCC) (Zhang et al., 2005). This could have resulted in a biased 'spike' appearing in the genetic data, creating a potentially misleading mutation hotspot in the COSMIC database (Figures 3.2 and 3.6a). A second possibility is that activation by this mutation could be tissue specific or context dependent; this work has not tested the effects of this mutation in epithelial cells similar to the ones in which the G697C mutations were found. However, another study that analysed a similar number of OSCC tumours from a different population did not detect any G697C mutations in FGFR3 (Aubertin et al., 2007). This supports the first interpretation of the data suggesting that population bias has created this mutation hotpot in FGFR3, reinforcing the importance of experimental testing during functional assessment of genetic data.

In summary, further experimental testing of the effects of these mutations (including in more relevant cancer-associated cell lines where available) is needed to understand the roles of these mutations in an oncogenic context. Such work can be guided by the experimental results shown here which suggest that mutations might impact several aspects of downstream signalling compared to WT, in addition to the direct and diverse alterations of KD activity and stability observed here for several activating cancer-associated mutations.

3.4 Discussion

In this chapter a comprehensive panel of cancer-associated mutations in the FGFR KD was screened *in vitro* in search of potential oncogenic drivers. A subset of mutations identified as genetic mutation hotspots and/or potential activating mutations from *in vitro* activity assays was further analysed in a cellular setting to assess the mutation-induced impacts on cellular transformation and signalling. The importance of two known driver mutations in FGFRs (N540K and K650E in FGFR3) has been reinforced, with N540K (and N540S) being analysed for the first time in FGFR3 and in the context of cancer. This work therefore consolidates previous work that analysed the activation status of similar mutations at these positions in FGFRs, specifically N549T/H in FGFR2 (Chen et al., 2007b) and K659E/M/N/Q/T and K650E/M/N/Q/T in FGFR2 and FGFR3 (Chen et al., 2013b, Huang et al., 2013b). This work therefore contributes to the growing recognition of highly activating mutations at these two positions as important oncogenic drivers, and suggests that while several aspects of the mechanisms of activation at each position are similar, other aspects differ (both in terms of direct KD activation as well as cellular effects). It was shown that the mechanism of direct KD activation by N540K in particular acts by increasing ATP binding affinity, increasing catalytic rate and reducing stability of the KD compared to WT. These effects might coincide to increase conformational flexibility of the KD, permitting a dynamic shift in the conformational population to more activated states. This can occur in a graded manner depending on the replacement as suggested for different K659 mutations in FGFR2 (Chen et al., 2013b). Graded activation of FGFRs by different replacements at the same residue has been shown even by early studies using FGFR3-K650 mutations in the context of dysplasia (Naski et al., 1996, Webster et al., 1996), suggesting that different replacements even at the same residue determine the extent of kinase activation and severity of the pathogenic phenotype. In this work, it has been shown that different mutations at the same residue at several different positions in the FGFR3 KD (I538, N540, D641, K650 and R669) display similar graded activation, with some of these replacements being more activating than others. Specifically, mutations I538V, N540K, D641G, K650E and R669G were more activating than I538F, N540S, D641N, K650N and R669Q respectively. In a clinical context, such subtle differences might propagate to large impacts on the response of patients to targeted treatments (discussed below). This work suggests that the mutation status of a therapeutic target must be taken into account during patient selection in the clinic, to help in identifying patients with highly activating mutations to which tumours are more likely to be addicted.

It has also been shown that rare, non-hotspot mutations in terms of the genetic data (such as FGFR3-R669G) might be activating and contribute to pathogenesis. Conversely, it has been shown that all genetic hotspot mutations (such as FGFR3-G697C) are not necessarily activating and that intrinsic biases in the compilation of genetic data by databases such as COSMIC might lead to the misleading generation of mutation hotspots. Additionally, due to the combined use of automated literature mining and manual curation of mutations, COSMIC curates all types of cancer-associated mutations including *de novo* cancer mutations (the majority) and acquired resistance mutations, with no annotations to distinguish between these other than reading the referenced literature sources. One example is the FGFR3 gatekeeper mutation V555M, a resistance mutation that was found to be activating in this work and in a recent study (Bunney et al., 2015) which might be mistaken as a *de novo* cancer mutation through COSMIC. Genetic data curated in the COSMIC database also has several other limitations. Firstly, to date, the mutation dataset is still curated from a relatively small number of samples, including both benign and malignant tumours, and is too small to be used for accurately inferring functional relevance (although it is constantly growing and is still an invaluable resource). Therefore, the absolute number of mutations in cancer is likely to be largely underestimated in COSMIC. Secondly, there is an intrinsic bias in that several studies from which mutation data is curated have only sequenced part of the genome, sometimes even only part of a gene, and sometimes without confirming the somatic status of the mutation. A variety of different screening methods are used, tailored to the type of mutation being searched for, for example insertions or deletions (such methods may miss the identification of missense substitutions, for example). Often these sequencing studies are only performed on regions of genes (particular exons) that are already known to harbour pathogenic mutations from previous work, from either cancers or congenital diseases, or both. In FGFR3, this has led to an explosion in the number of genetically identified mutations creating possibly overrepresented hotspots at positions such as S249C and K650E due to the relatively early identification of mutations in the D2-D3 linker and the A-loop as crucial for pathogenesis in cancer and dysplasia (sections 3.1.4 and 3.1.5). Conversely, a region such as the αEF-β12 loop in the KD (for example) has not been previously associated with cancer, so mutations in this region (for example R669G) may be underrepresented in databases such as COSMIC as sequencing studies tend to ignore these regions when searching for mutations in new samples. This work attempted to compile a comprehensive and representative panel of FGFR mutations from all regions of the KD using a semi-biased selection process, incorporating mutations irrespective of the number of observations or position in the KD. It was found that both hotspot and non-hotspot mutations can be activating and might play roles in oncogenesis and importantly, no trend was observed between the number of observations in cancer

and activation status. Therefore, this work suggests that functional validation of the genetic data in databases like COSMIC must be performed, as both hotspot and non-hotspot genetically-curated mutations can be important for pathogenesis. An interesting observation was that the three most activating mutations in this work (N540K, K650E and R669G) are also commonly observed in dysplasias in the syndromes with the worst prognoses (Wilkie, 2005). This suggests that mutations that co-occur in both congenital diseases and malignancies might aid in the validation of possible cancer-associated driver mutations based on their links to the pathogenesis of both types of proliferative diseases. The mechanisms of action of such mutations proposed in this work might therefore apply to both congenital diseases and cancers.

All pathogenic mutations found in the cytoplasmic region of FGFR in the cancer mutation databases (including COSMIC, TCGA and ICGC) were also analysed for their functional impacts using a combination of computational algorithms (Patani et al., 2016). This work showed that while many predictive algorithms are relatively advanced at identifying pathogenic loss-of-function mutations, *in silico* predictions of activity, stability and pathogenicity do not yet accurately represent the experimental reality when it comes to gain-of-function mutations in FGFRs, such as those important for the development and maintenance of cancers and other congenital disorders. Additionally, the computational tools performed better when attempting to link structure and function for FGFR residues that have a well-defined function or are found conserved across kinases, but tended to inherently overlook infrequently occurring variants or those that are not conserved across FGFRs or kinases. Therefore, although *in silico* predictions of function are an invaluable and complementary resource, this work emphasises the importance of experimental evaluation in linking genetic data to function. This is also important during selection of targets for drug discovery, to ensure critical selection based on pre-clinical data-driven predictions of the oncogene addiction status of a kinase or its mutant. In general, the use of a combination of complementary techniques to biochemically determine the roles of pathogenic mutations in cancer in this work presents a combinatorial platform for similar studies in other RTKs and EPKs involved in cancer. The data generated from such studies can then be used to better train predictive algorithms to improve the quality of future *in silico* predictions of the effects of mutations on function, particularly gain-of-function.

Useful functional information can be gained from assessing the direct impacts of mutations on the activity of isolated KDs in solution, as shown in this study and by others (discussed above). In particular, this work used several variations of *in vitro* kinase activity assays to comprehensively assess the impact of mutations on the activation

status and kinetics of catalysis by KDs in solution, something which is much more difficult to investigate in the context of the full length receptor in cells. However, such minimalist experiments also have some limitations. For example, due to the reductionist nature of the *in vitro* experimental system used in this work, it might not be possible to accurately assess the functional impact of mutations that affect kinase activity by mechanisms other than direct activation of isolated KDs. For example, mutations such as E466K that lie within the intracellular juxtamembrane region (a region implicated as vital for activation of full length RTKs such as EGFRs and perhaps FGFRs, see section 1.5.3) which has been shown to greatly reduce the thermal stability of the isolated FGFR3 KD (Figure 3.15), might only exhibit elevated kinase activity in the context of the full length receptor comprising an intact, anchored juxtamembrane segment. This mode of kinase activation might be missed by the experimental design used in this work. Although efforts were made to assess the impact of a subset of mutations in the context of full length FGFR3 in cells using a simple cellular model (Figures 3.16-3.20) more work needs to be done to elucidate their functional roles in cells. In this work, the focus was on assessing the mechanistic impacts of highly activating mutations identified from the *in vitro* activity assays. In particular, the effects of mutations in the correct cellular context need to be addressed, guided by the cancer cell types in which the mutations were originally identified (see Table 3.2 for examples). The addiction status of cancers to these mutations also needs to be determined to validate specific variants as individual therapeutic targets. Combining *in silico* molecular dynamics simulations with structural studies using full length mutant FGFRs will provide clarification of the modes of FGFR activation adopted by different mutations. Newly identified gain-of-function mutations from such studies can then be rigorously tested in cells and *in vivo*, including detailed analyses of changes in downstream signalling. Experiments to test these hypotheses are currently in progress in the lab for activating mutations and also for the three most destabilising mutations identified in this work, namely E466K, I538F and N540K.

The correct cellular context might be vital for some mutations to exert their cell-transforming effects, as was proposed for the R669G and G697C mutations investigated in this work in cells. This work showed that highly activating mutations K650E and N540K were activating both *in vitro* and in cells, and transformed NIH 3T3 cells morphologically as well as by promoting anchorage independent growth. However, despite R669G being highly activating in *in vitro* assays and even showing activation in cells, this mutation did not morphologically transform NIH 3T3 cells or enable anchorage independent growth. Similarly, G697C was not activating *in vitro* or in cells despite a previous study suggesting it is an activating mutation when expressed in an insect cell line (Zhang et al., 2005). In addition, Y647C which was reported to act as a mechanism of acquired

resistance to EGFR targeting by activating FGFR signalling in a pre-clinical human tumour model (Crystal et al., 2014) was not found to be activating in this work and did not morphologically transform NIH 3T3 cells. Although very useful for providing a direct and comparative readout of cellular transformation between mutants, there are several limitations of NIH 3T3 cells used to generate the model stable cell lines in this work. Firstly, the cells are of mouse origin and therefore have subtle overall differences compared to human cells. Secondly, the cells are not of epithelial origin, as are the majority of tumours in which FGFR cancer mutations are detected. Finally, generation of stable cell lines and subsequent overexpression of FGFRs might not accurately represent the situation in cancer and so may result in slight differences in cell morphology, signalling and anchorage independent growth compared with a real tumoural context. Therefore, the mutations need to be assessed for their functional impact in cell lines of epithelial origin as well, similar to the cancer cells in which the mutations were identified. In addition, this could be done using three-dimensional tissue culture models in which the tumour microenvironment can be better represented. These combinations of experiments will help to ascertain whether the cell transforming and/or oncogenic effects of mutations will present in the correct cellular context or tumour microenvironment.

Interestingly, it has recently been proposed that FGFR3 mutations such as S249C and K650E might only impart highly oncogenic effects *in vivo* when found in combination with other deregulated genes (such as loss of the tumour suppressors p53 or pRB) in specific cellular contexts in urothelial cells (Zhou et al., 2016). This can be further complicated by the fact that tumours often comprise heterogeneous clonal populations of cells that arise from different aberrations and propagate to different tumour grades within the same cancer, including urothelial tumours harbouring FGFR3 aberrations (Tomlinson et al., 2007a). This work has also shown preliminary data suggesting FGFR mutations increase crosstalk with different intracellular signalling pathways compared to WT (Figure 3.20). It was also noticed that two independent K650E cell lines differently activated downstream signalling pathways via either pERK or pAKT (Figure 3.19c-d). Although experimental validation of these initial findings is required, these data suggest that oncogenic signalling networks induced by FGFR mutations can be multifaceted. They might even include co-activation of other RTKs as both *de novo* mechanisms of activation by aberrant FGFRs and/or in response to targeted inhibition of other RTKs or FGFRs (as discussed in section 3.1.6). Therefore, the cellular context, including the mutation and co-activation status of other genes, might be a crucial determinant for FGFR mutations to act as oncogenic drivers in cancer. This remains to be rigorously investigated for the majority of FGFR cancer-associated mutations.

This work performed the first kinetic analysis of a large panel of FGFR variants including calculations of ATP affinity constants and rates of reaction. It was shown that N540K directly impacted KD activity by increasing its rate of phosphorylation, at least partly due to its increased affinity for ATP (Figure 3.14) and possibly partly due to reduced KD stability, as discussed above. However, in many of these assays, the initial ATP concentrations used were below the $K_{M(ATP)}$ of the KDs being investigated, and reaction time was longer than that required for maintaining linearity. In autophosphorylation and some substrate phosphorylation experiments (Figures 3.9b, 3.10a-b, 3.11c and 3.12b-c) which were assays to compare variant activity, this helped to maximise the observed differences in activity between variants to clearly rank them in order of activation status compared to FGFR3-2C. Here, the relatively low 50 μM ATP and longer 45 minute reaction time used were of less importance than the fact that these values were identical for each variant in the same assay; however they may have contributed to the 'levelling off' of curves, particularly by more activating variants, as the reactions slowed down due to consumption of large proportions of substrates. An exception is the substrate phosphorylation experiment (Figure 3.11a-b) which was more than just a comparative screening experiment as K_D values were also extracted from the fitted curves, but this assay was performed with at least a 2-3 fold excess of ATP (compared to the $K_{M(ATP)}$ of variants being assayed, Figure 3.14d) and with a 10 minute reaction time, so that it sufficiently conformed to the laws of steady-state enzyme kinetics. Time course and substrate (ATP or polyE_4Y_1 peptide) titration experiments (Figure 3.9c-e and 3.14) were also performed with at least a 2-3 fold excess of ATP and with a 10 minute reaction time to ensure signal linearity. In addition, the non-linear regression analyses of these data performed using equations built-in to the Prism software (equations **1**-**4**, section 2.9) are designed for assays with these types of experimental set-up. The assumptions involved in these analyses are already listed in section 2.9.

Interestingly, despite the general trends observed in this work from the *in vitro* kinase assays (phosphorylation responses and reaction rates increasing with more activating variants), there are a few key instances when these trends do not hold true. For example, the maximum extents of phosphorylation observed in Figure 3.14a and 3.14c do not appear to mimic the ranked activation status of the variants, and in addition the reactions differently slow down over time. In particular, reactions involving N540K have a very high initial rate but then appear to 'level off' very rapidly compared to all other variants, while reactions involving R669G/Q (Figure 3.14a) and K650N (Figure 3.14c, left panel) appear to barely decelerate within the tested duration or ATP concentration range respectively. These reaction properties are likely to emerge as a consequence of combinations of possible phenomena including (but not limited to) concentration threshold dependent

product and/or substrate inhibition, and degradation of KD proteins over the course of reactions (the N540K KD is a particularly unstable activating variant, see Figure 3.15). These will probably result from (small) mutation induced differences in KD structure and chemical microenvironments that affect substrate binding (and release), allosteric sites, and overall KD function. This presents as a range of mutation specific *in vitro* properties as described. However, in the context of a cell where substrates and products are in constant flux, different permissive kinetic properties might allow distinct extents and rates of reactions to those observed in the *in vitro* assays in this work.

Another limitation of the *in vitro* kinase assay arises because nanomolar concentrations of kinases are required to ensure strict signal linearity and to overcome 'assay wall' (phenomenon by which the lowest IC_{50} value that can be measured by the assay is half the initial kinase concentration) when developing the assay for the determination of catalytic and inhibitory binding constants for potent inhibitors (Yli-Kauhaluoma and Tuominen, 2011), as was a future aim of this assay. Such low concentrations cannot be measured by standard laboratory instrumentation and therefore for kinetic calculations involving absolute kinase concentrations (such as determination of k_{cat} and $k_{cat}/K_{M(ATP)}$) a theoretical error (10%) was applied to the calculated kinase concentrations. This was based on the expected errors involved in measurement of the concentrated protein stock solutions and subsequent dilution for assays. Calculations of affinity and rate constants were then performed both with and without this theoretical error, and resulted in only small differences in the propagated errors. Subsequently, Figure 3.14e and Table 3.4 were created using the resulting calculated errors for k_{cat} and $k_{cat}/K_{M(ATP)}$. Although these analyses were used to define several kinetic parameters and rate constants using a very robust activity assay, all parameters should ideally be calculated using measurements from at least two different types of assays to ensure the absolute values measured are accurate. However, it should be noted that within a cell, kinase activity probably does not actually follow Michaelis-Menten kinetics as protein substrates are unlikely to be in excess in the crowded intracellular environment. Therefore, the values measured in this work in direct and comparable assays in parallel provide valuable insights into the relative catalytic rates and substrate binding affinities compared to WT, which are useful for informing future FGFR-targeted drug discovery. However, these data provide little information about the kinetics of catalysis by different mutations within the cell; single-molecule FRET experiments (for example) can be useful for investigating such properties further, including in a cellular context or using saturating ATP concentrations to provide valuable complementary information to that produced in this work.

Highly activating mutations were found to activate downstream signalling compared to WT. It was also noticed that N540K and K650E resulted in higher receptor expression levels compared to WT and compared to other less activating mutations (and even compared to highly activating R669G). It was not known whether this was a feed-forward mutation-induced mechanism that is involved in oncogenesis, or whether it is a consequence of artefacts incurred during the production of (non-clonal) stable cell lines. Increased receptor expression compared to WT has previously also been observed for stable cell lines harbouring activating FGFR2 mutations (Byron et al., 2013). Additionally, a similar association has been previously observed in a study using urothelial cancer samples which showed that 85% of tumour samples harbouring FGFR3 mutations also overexpressed the receptor (Tomlinson et al., 2007a). This was observed across urothelial tumours of all grades and stages, albeit in a greater proportion of lower grade tumours. In this work, it was suggested that transformation of NIH 3T3 cells is independent of the expression level of the WT or mutant receptors. However, it remains to be investigated whether threshold levels of mutant FGFRs are required for the alterations in downstream signalling observed in this work. Additionally, it is currently less clear how mutations found to be moderately activating (such as I538V, N540S, V555M, D641G/N, K650N and R669Q from this work) might be involved in oncogenesis, as none of these variants caused transformation of NIH 3T3 cells. In particular, FGFR3-D641 mutations did not transform NIH 3T3 cells in this work, despite the activating hotspot mutation D816V at the corresponding position in the KIT receptor having been shown to transform the mast cells in which it exerts its oncogenic effects (Hashimoto et al., 1996) and KIT-D816V also confers resistance to the inhibitor imatinib (Frost et al., 2002). The V555M gatekeeper mutation has been previously shown to provide acquired resistance to several ATP-competitive inhibitors (section 3.1.6) and this work suggests that this might occur through steric hindrance of these inhibitors binding in the ATP (and thus inhibitor) binding pocket as suggested by the increase in ATP binding affinity of the FGFR3-V555M KD. Several of these moderately activating mutations increased the catalytic efficiency and ATP binding affinity of the FGFR3 KD by varying degrees (Figure 3.14d-e). This might 'prime' the mutant FGFRs by increasing basal ligand-independent activation so that they may present as oncogenic drivers in the optimal cellular context (discussed above). In addition, the effects of double mutations in FGFRs have not been analysed in this work, but double mutations have been previously observed in cancer including in FGFR3 in urothelial carcinoma (Tomlinson et al., 2007a). Their effects have only recently begun to be studied in FGFRs, such as in a recent structural analysis of double KD mutations in FGFR2 showing that double mutations in different parts of the KD confer enhanced, additive activation compared to their single-mutant counterparts (Chen et al., 2017). Although the occurrence of double mutations in the same allele is

predicted to be extremely rare in cancer, a double FGFR3 mutation including N540K and another KD mutation on the same allele has been observed in a dysplasia where it increase the severity of the disease phenotype (Pannier et al., 2009). Further analysis of these mutations, including in cells (using the relevant cancer cell lines where available) will provide information such that even malignancies and congenital disorders with rare mutations can be treated in a patient-personalised manner in the future.

The clinical efficacy of FGFR targeted therapies has so far fallen short of the expected success of these drug candidates despite their success in pre-clinical testing (Babina and Turner, 2017). This could be due to the fact that patients have so far not been selected for clinical trials based on mutation, and importantly addiction, status of their tumours to aberrant FGFR signalling, possibly resulting in suboptimal clinical success of several pre-clinically potent FGFR inhibitors. Additionally, the efficacy of pre-clinical drug candidates is normally only experimentally verified using WT protein targets, with only rare functional assessment of the impact of mutations on inhibitor efficacy. In the case of kinases, this can be particularly relevant for mutations in the KD which can directly affect the activity and conformation of the KD and/or inhibitor binding site. To test this, the *in vitro* kinase assay developed here was used to assess the impact of several FGFR activating mutations identified in this work on the efficacy of a panel of FGFR inhibitors currently in clinical trials (Patani et al., 2016). This work clearly demonstrated that different mutations affect drug efficacy to differing degrees. Rarely, mutations increased the efficacy of specific inhibitors, while others (more commonly) reduced efficacy. Impacts on potency differed for different binary combinations of mutations and inhibitors, with no trend that could be predicted without experimental testing. For example, the V555M gatekeeper mutation and N540S (also in the ATP binding pocket) conferred particular resistance to the FGFR-selective inhibitor JNJ-42756493 while a different replacement at the same 'molecular brake' residue to N540K provided most resistance to another FGFR-selective inhibitor, AZD4547, with less impact on JNJ-42756493 binding. Other mutations in the vicinity of the ATP binding pocket such as moderately activating I538V also conferred resistance to these two inhibitors as well as the multi-kinase inhibitor Dovitinib. In contrast, activating mutations distant from the ATP (and inhibitor) binding site (K650E and R669G) did not greatly impact inhibitor binding. None of the mutations tested conferred a change in the sensitivity to the multi-kinase inhibitor Ponatinib. This agrees with a recent study which showed that the corresponding N549K mutation in FGFR2 exhibited increased resistance to the inhibitors Dovitinib and PD173074, but not Ponatinib, compared to FGFR2-WT in an endometrial carcinoma cell line (Byron et al., 2013). Such subtle differences in the primary sequence leading to conformational changes in the KD and alteration of the chemical environment in the

inhibitor binding site can therefore lead to large impacts on the success of targeted inhibition, including in the clinic. Resistance to targeted therapies is one of the key challenges still to be anticipated and overcome in the clinic. Therefore, a combination of functional studies such as those performed in this work, combined with a pre-clinical assessment of the impact of mutations on inhibitor potency, should be a crucial step in data-driven targeted therapeutics in the future and similar principles can be applied across RTK and EPK targets in cancer. The true success of clinical targeted therapies, particularly ATP-competitive inhibitors, might depend on such multifaceted information. Specifically, the impacts of subtle point mutations on the efficacy of clinical candidate drugs have not always been taken into account during clinical trials (although this is rapidly changing). This might have led to sub-optimal patient selection which could be at least partly responsible for the unexpectedly low clinical response rates observed to date. Pre-clinical studies that take into account the effects of mutations should help to boost the clinical success of targeted therapies which are currently less successful at targeting FGFRs harbouring point mutations compared to their relative success against cancers with *FGFR* amplifications and fusions which contain receptors with a WT sequence (Babina and Turner, 2017). This should help to combat clinical resistance to FGFR and other RTK inhibitors through knowledge of the mechanisms of action of different pathogenic mutations against them. In addition to resistance through specific point mutations in FGFRs, other mechanisms of resistance can develop in tumour cells through co-activation of other RTKs or other downstream signalling pathways, bypassing the need for FGFRs to maintain a robust oncogenic signal. As previously mentioned, this work showed preliminary results demonstrating crosstalk between FGFR signalling and other downstream signalling networks such as the Wnt signalling pathway (Figure 3.20). This pilot experiment also showed that pathogenic mutations might affect these signalling patterns, possibly by altering the networks in favour of maintaining the robust oncogenic signal and/or to diversify the oncogenic outcomes. Treating tumours with FGFR inhibitors can therefore also lead to oncogene switching such that the tumour becomes addicted to signalling via another kinase, simultaneously becoming resistant to FGFR inhibition through loss of oncogene-addiction to FGFR signalling due to the targeted therapy. Therefore, knowledge of the combined mechanisms of resistance through FGFR mutation, oncogene switching and/or co-activation of oncoproteins is required to target cancers with appropriate combination therapies to robustly eliminate them. Further signalling level data combined with studies as in this work are therefore required to inform a truly personalised approach to FGFR-targeting in the future of anti-cancer therapeutics.

Chapter 4: Characterisation of endocrine FGFR signalling complexes

4.1 Introduction

The endocrine sub-axes of FGFR signalling refer to the more recently discovered, and less well-researched, modes of signalling that have key regulatory differences compared to 'canonical' paracrine FGFR signalling. Endocrine signal transduction requires the constitutive recruitment of a member of the Klotho family of FGFR co-receptors into the ternary signalling complex instead of the HS co-factors required for paracrine signalling. Inclusion of Klotho proteins, exclusion of HS and switching of FGF ligands in the ternary signalling complexes from paracrine to endocrine sub-family members (FGF19, FGF21 and FGF23 in humans) all serve to alter the mode of FGFR signal transduction from paracrine to endocrine (Belov and Mohammadi, 2013, Goetz and Mohammadi, 2013). This leads to changes in the physiological outcomes of signalling from controlling development, growth, migration and tissue homeostasis to hormonally regulating diverse metabolic processes including glucose uptake, bile acid synthesis and phosphate and vitamin D homeostasis. Strong tissue specific expression of the Klotho family proteins coupled with selective inclusion of specific endocrine FGFs and FGFR isoforms in the ternary complexes dictate the metabolic outcome of the transduced signal (Kurosu and Kuro, 2009).

The important discovery that endocrine FGFR signalling mediates several aspects of normal metabolism coincided with the growing observations that the most common side-effects in response to targeted FGFR therapeutics (Table 1.4) in pre-clinical tumour models were due to on-target inhibition of specific endocrine FGFR signalling pathways, most commonly involving FGF23 (Dieci et al., 2013, Brooks et al., 2012). Additionally, in contrast to FGFR point mutations involved in cancer discussed in Chapter 3, an aspect of aberrant endocrine FGFR signalling such as amplification of *FGF19* has also been identified as a *de novo* driver of oncogenesis in specific cancers (Sawey et al., 2011). The same *FGF19* aberration has also been shown to confer sensitivity to targeted FGFR inhibitors in specific subsets of cancer cell lines, but only when the cognate Klotho family co-receptor required for FGF19 signalling is concurrently expressed (Guagnano et al., 2012). Finally, FGF21 has been identified as an important modulator in several clinically important metabolic diseases such as diabetes and obesity. These involvements of components of endocrine signalling in oncogenesis, modulation of toxicity profiles in response to targeted FGFR therapies and in widespread metabolic diseases provides a strong rationale to further study the relatively unknown sub-axes of endocrine FGFR signalling further.

In this Chapter, the three components of the endocrine ternary signalling complexes will be introduced from a structural and biochemical perspective, with a focus on the Klotho family of co-receptors and the endocrine FGF19 sub-family proteins (the FGFR family of RTKs has already been described in detail in Chapter 1). In particular, the focus will be on βKlotho, FGF19 and FGF21 and the roles of their resultant ternary FGFR complexes in endocrine signalling. The importance of composition of the signalling complexes and tissue specificity of expression on context-dependent metabolic effects will be presented. The roles of deregulated endocrine signalling on various pathologies, including cancer, will be discussed. Finally, the development of components of the ternary complexes and related biomolecules as therapeutics for several diseases, including cancer, will be reported.

4.1.1 The Klotho family of FGFR co-receptors

The three members of the Klotho family of proteins are αKlotho (alias KLA), βKlotho (alias KLB) and γKlotho (alias lactase-like protein, LCTL). The *Klotho* gene (for the αKlotho protein) was first discovered as an anti-ageing gene in a knockout mouse model that displayed signs of human-like ageing and reduced lifespan (Kuro-o et al., 1997). Later, the mouse *βKlotho* gene was also identified based on sequence identity of the protein to *Klotho* of 44% (Ito et al., 2000). This was followed by discovery of *γKlotho*, the third Klotho family member identified to date (Ito et al., 2002). The three proteins make up a sub-family within the superfamily of type I β-glucosidases based on their sequence similarities to glycosyl hydrolases (Henrissat and Bairoch, 1993). Most research has so far focussed on the anti-ageing αKlotho, with emerging interest in the βKlotho branch of signalling. Very little is known about γKlotho and its functions so this Klotho family member will not be discussed in much detail in this work except for the rare cases where data is available.

4.1.2 The endocrine sub-family of FGF ligands

Members of the FGF19 sub-family of FGF ligands comprise FGF19 (mouse orthologue is FGF15), FGF21 and FGF23 (see section 1.5.1 and Figure 4.1). Each FGF forms complexes with specific Klotho family members and FGFRs and this dictates the outcomes of endocrine signal transduction. The first endocrine FGF to be identified was FGF19, which was discovered to bear the least sequence homology to all human FGFs known at the time (Xie et al., 1999, Nishimura et al., 1999). It was found to be the orthologue of mouse FGF15 discovered previously due to 53% sequence identity between the two proteins (McWhirter et al., 1997). FGF21 was discovered next in mouse embryos (Nishimura et al., 2000), followed soon after by FGF23 (ADHR Consortium, 2000, Yamashita et al., 2000). All were predicted to be secreted proteins and were

grouped into their own sub-family of endocrine FGFs (Itoh and Ornitz, 2008, Itoh and Ornitz, 2011). Compared to paracrine FGFs, they were all found to require specific Klotho co-receptors for high affinity binding to and signalling through FGFRs in a HS independent manner. They comprise unique structural features that facilitate their signalling capabilities as hormones compared to other FGF sub-families, as mentioned in section 1.5.1 and further described in sections 4.1.9 and 4.1.10 below.

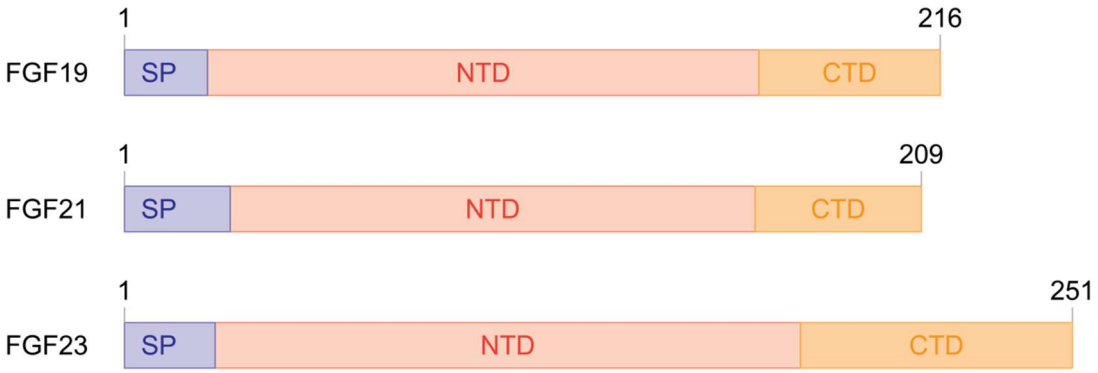


Figure 4.1 Domain organisation of the human FGF19 subfamily of endocrine FGF ligands. Schematic representation of the domain structures of endocrine FGFs (to relative scale). All three FGFs comprise a short N-terminal secretion signal peptide (SP), a 'core' N-terminal domain (NTD) which is highly conserved among all FGFs and a C-terminal domain (CTD) referred to as the C-terminal tail, which confers Klotho protein binding specificity. The amino acid length of each human FGF protein is indicated. See also Figures 1.2 and 1.6.

4.1.3 FGFR isoforms

FGFRs have already been introduced in Chapter 1. All FGFR isoforms (FGFR1-4) can form endocrine signalling complexes, but several studies have shown that it is particularly the canonical c, and not the alternative b, splice isoforms of FGFR1-3 that bind and transduce signals in complexes with Klotho proteins and endocrine FGFs (Xie et al., 1999, Urakawa et al., 2006, Kurosu et al., 2007, Ogawa et al., 2007). Some ternary complex combinations can form *in vitro* and in cellular models but cannot transduce a signal; therefore few specific ternary complexes appear to be functionally relevant *in vivo* and this is discussed in more detail in sections 4.1.9 and 4.1.10 below.

4.1.4 Domain organisation and possible functions of Klotho proteins

The Klotho proteins are transmembrane glycoproteins with a very short cytoplasmic domain, a single-pass type I transmembrane helix and a large ectodomain with variable numbers of putative *N*-glycosylation sites (Table 4.1). The ectodomain of αKlotho and βKlotho comprises two β-glucosidase-like tandem repeat domains termed KL1 and KL2 separated by a short linker, while γKlotho has only one KL domain (Figure 4.2). To date, no atomic structural information exists for any Klotho proteins or any of their individual domains, with or without FGF or FGFR binding partners. This information is eagerly sought in order to understand the molecular basis for the crucial roles of Klotho co-receptors in the switch from paracrine to endocrine signalling. Additionally, structural information will help to decipher the stoichiometry of signalling complexes and will aid in designing therapeutics targeting these ternary complexes for the treatment of a range of metabolic pathologies and specific malignancies in which Klotho-dependent signalling is deregulated (reported below), as well as in combating the adverse effects observed in patients in response to clinical FGFR targeting (reported in section 4.1.18).

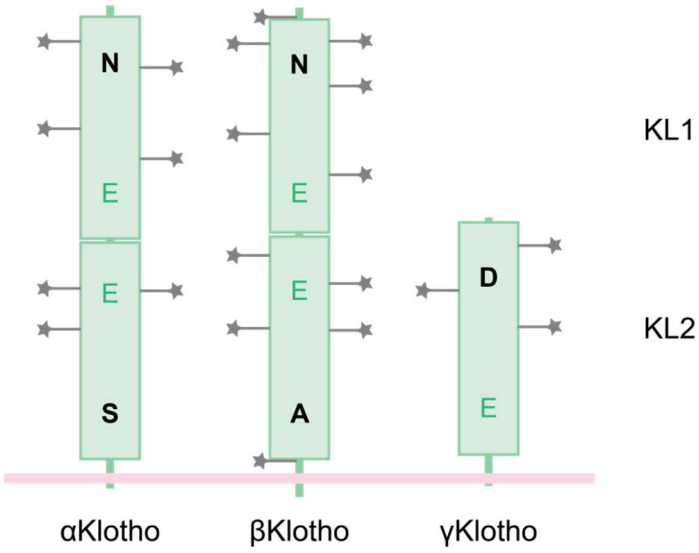


Figure 4.2 Domain organisation of the human Klotho family of transmembrane co-receptor proteins. The three Klotho family members known to date (αKlotho, βKlotho and γKlotho), based on sequence homology, are schematically illustrated (to relative scale). They all comprise an extracellular region, a single pass type I transmembrane helix and a very short C-terminal cytoplasmic region. The large ectodomains of αKlotho and βKlotho (almost 1000 amino acids in length) comprise two β-glucosidase-like tandem repeat domains termed KL1 and KL2 (labelled) separated by a short linker, while the ectodomain of γKlotho comprises only one KL domain. The entire extracellular region comprising KL1-KL2 (αKlotho and βKlotho) or just the individual KL domains for all three Klotho proteins are also shed into the circulation where they are referred to as soluble Klotho. At least one of the essential catalytic aspartate (D) residues required for β-glucosidase activity is missing in all the KL domains, and the replacement residues (N, A, S or E) have been labelled on each KL domain. For this reason, the KL domains are predicted to be catalytically inactive. Putative *N*-glycosylation sites on the ectodomains of all three Klotho proteins have been labelled with grey stick-and-star markings (positions to scale).

Table 4.1 Human Klotho family members predicted protein and domain lengths, protein molecular masses and numbers of putative glycosylation sites.

	αKlotho	βKlotho	γKlotho
Protein length (amino acids)	1012	1014	567
Extracellular domain (amino acids)	948	996	519
TM domain (amino acids)	21	21	21
Intracellular domain (amino acids)	10	27	5
Non-glycosylated molecular mass (kDa)	116	120	65
Glycosyl hydrolase (KL) domains	KL1 (57-506) KL2 (515-953)	KL1 (509-516) KL2 (517-967)	KL (33-503)
Putative glycosylation sites (number)	7	11	3

4.1.5 Possible enzymatic functions of KL domains

Despite their sequence similarity to β-glucosidases, all KL domains appear to lack an essential catalytic aspartate residue required for glucosidase activity (Figure 4.2) and are therefore predicted to be catalytically inactive (Ito et al., 2002). A few studies have suggested that αKlotho might exhibit weak β-glucuronidase activity (Tohyama et al., 2004, Chang et al., 2005) while others have suggested that sialidase activity of αKlotho might regulate ion channel activation by specifically acting on α2,6-linked sialic acids (Cha et al., 2008). To date, no such studies have implicated βKlotho or either of its individual KL domains to perform any glucosidase, glucuronidase or sialidase activities. Additionally, whether or not either of the KL domains in any Klotho family member retains any of these enzymatic activities is still an open, unresolved question. However, it is hypothesised that the KL domains may have retained the ability to bind to sugars such as galactosides (Kilkenny and Rocheleau, 2016). This might have important implications for cell surface signalling with Klotho proteins possibly retaining the ability to bind, but not metabolise, cell-surface glycan moieties (discussed further below).

4.1.6 Roles of Klotho ectodomains as secreted humoral factors

The ectodomain of αKlotho is also shed into the circulation as a hormone, particularly in humans and to a lesser extent in mice (Matsumura et al., 1998, Shiraki-Iida et al., 1998). In humans it is thought to originate predominantly from the kidney, the organ where full length membrane-bound αKlotho is most highly expressed (Olauson et al., 2017), as described in section 4.1.7 below. The secreted ectodomain has been named soluble Klotho (sKlotho) and refers to the entire ectodomain comprising both KL domains (Figure 4.2), although secretion of individual KL domains has also been observed (Chen et al.,

2007a). However, the human αKlotho KL1 domain, compared to the entire ectodomain, could not efficiently induce FGF23 signalling in cells or hypophosphatemia in mice (Abramovitz et al., 2011) suggesting that the ectodomain is needed in its entirety for the hormonal functions of αKlotho. sKlotho has been detected in bodily fluids such as sera and cerebrospinal fluid where it is released post-translationally through enzymatic cleavage from membrane-bound full length αKlotho, rather than being directly transcribed (Imura et al., 2004). Cleavage is catalysed at least in part in an insulin dependent manner by A Disintegrin and Metalloproteinase (ADAM) family sheddases ADAM10 and ADAM17 (Chen et al., 2007a) and possibly by other secretase proteins (Bloch et al., 2009) *in vivo*, although sKlotho might also result as a product of alternative mRNA splicing (Hu et al., 2010, Mizuno et al., 2001). Physiological levels of sKlotho appear to vary between distinct genetic populations (Schmid et al., 2013) and have been shown to decline with age (Yamazaki et al., 2010). The levels of protein are important in regulating a range of pathologies, and can serve as useful biomarkers in disease (see section 4.1.19).

4.1.7 Expression profiles of Klotho, FGFs and FGFRs: implications for signalling

It has been shown that the Klotho proteins are expressed in a limited range of tissues with discrete and specific expression for each family member. In adult mice, αKlotho is expressed predominantly in the kidney and in the brain (Kuro-o et al., 1997, Chang et al., 2005, Fon Tacer et al., 2010), with lower levels of expression in other organs such as the thyroid and arterial vasculature, and similar expression patterns are observed in humans (Lim et al., 2015). The γKlotho protein is expressed highly in the eye, brown adipose tissue (BAT), skin and kidney (Ito et al., 2002, Fon Tacer et al., 2010, Zhang et al., 2017b). Expression of βKlotho is more widespread that the other Klotho family members, with highest levels of expression observed in the liver, pancreas, gall bladder and BAT and white adipose tissues (WAT) with lower levels in the gut, skin, arterial vasculature, reproductive tissues, skeletal muscle and hypothalamus (Ito et al., 2000, Lin et al., 2007, Fon Tacer et al., 2010, Adams et al., 2012, Bookout et al., 2013, Ding et al., 2012). It has emerged that the specificity of endocrine signalling depends on these tissue specific expression profiles of Klotho family members. In contrast, FGFRs are much more ubiquitously expressed across tissues compared to the Klotho proteins (Fon Tacer et al., 2010), and impart a much lower influence on the tissue specificity of signalling. Specific examples are discussed in sections 4.1.9 and 4.1.10.

The endocrine FGFs also exhibit discrete tissue specific expression profiles, however as mentioned previously, they are secreted hormones that exert their physiological effects mostly away from their sites of production. FGF19 (or FGF15 in mice) is expressed

mainly in the small intestine (Inagaki et al., 2005, Fon Tacer et al., 2010) but exerts its effects primarily on the liver and gall bladder where it dampens bile acid synthesis and gluconeogenesis, promotes post-prandial protein and glycogen synthesis and stimulates filling of the gall bladder (Inagaki et al., 2005, Choi et al., 2006, Lin et al., 2007, Kir et al., 2011, Potthoff et al., 2012). It has also been implicated to exert anti-diabetic and anti-obesity effects in mice (Tomlinson et al., 2002, Fu et al., 2004). FGF21 appears to have a broader tissue expression profile than FGF19 with sites of production including the liver, pancreas, testes, gut, thymus, and BAT (Nishimura et al., 2000, Fon Tacer et al., 2010) but circulating FGF21 is thought to be mostly liver derived and its expression is induced in response to fasted states in mice (Nishimura et al., 2000, Badman et al., 2007, Inagaki et al., 2007, Lundasen et al., 2007, Lee et al., 2011, Markan et al., 2014). In other tissues such as WAT, BAT and skeletal muscle FGF21 expression is induced in response to various other transcription factors involved in nutrient-sensing, reviewed in (Owen et al., 2015, Kilkenny and Rocheleau, 2016). The primary sites of action of FGF21 are the liver, pancreas and adipose tissues where it regulates a range of metabolic outcomes such as nutritional stress-induced energy homeostasis in liver and glucose uptake, insulin and leptin sensitisation, ketogenesis, lipolysis and thermogenesis in WAT and BAT, reviewed in (Owen et al., 2015, Itoh et al., 2015). Its primary role is thought to be stimulation of glucose uptake, first shown in 3T3-L1 adipocytes (Kharitonenkov et al., 2005). The potential anti-diabetic, anti-obesity and life extending effects of FGF21 have fostered great therapeutic interest in the hormone, as presented in section 4.1.19 below. The final endocrine FGF, FGF23, is expressed in bone, heart, thyroid, brain, lung, thymus, spleen, muscle and liver (ADHR Consortium, 2000, Liu et al., 2003, Fon Tacer et al., 2010). It targets the kidney and heart to exert its roles in phosphate and vitamin D homeostasis (Shimada et al., 2004). The predominant sites of endocrine FGF expression and metabolic action are summarised in Figure 4.3. Of the three endocrine FGFs, multiple studies have revealed that FGF19 and FGF23 also affect proliferation and have mitogenic activities in addition to their metabolic roles, while FGF21 uniquely has only metabolic activities without any proliferative or oncogenic effects (discussed further below).

It was noticed that there were similarities between phenotypes of multiple knockout mouse models for different combinations of αKlotho, βKlotho, FGF19, FGF21, FGF23 and various FGFRs. This led to the hypothesis that specific Klotho co-receptors were required for binding to specific combinations of FGFs and FGFRs for signal transduction. A wealth of further experimentation in mice and *in vitro* biochemical experiments were employed to decipher the roles of the proteins in relation to each other, leading to the discovery of the FGF-FGFR-Klotho ternary complexes and the requirements for Klotho

co-receptors for endocrine signalling. The biochemical evidence supporting the functional relevance of the specific complexes thought to be important to date is presented in the next section.

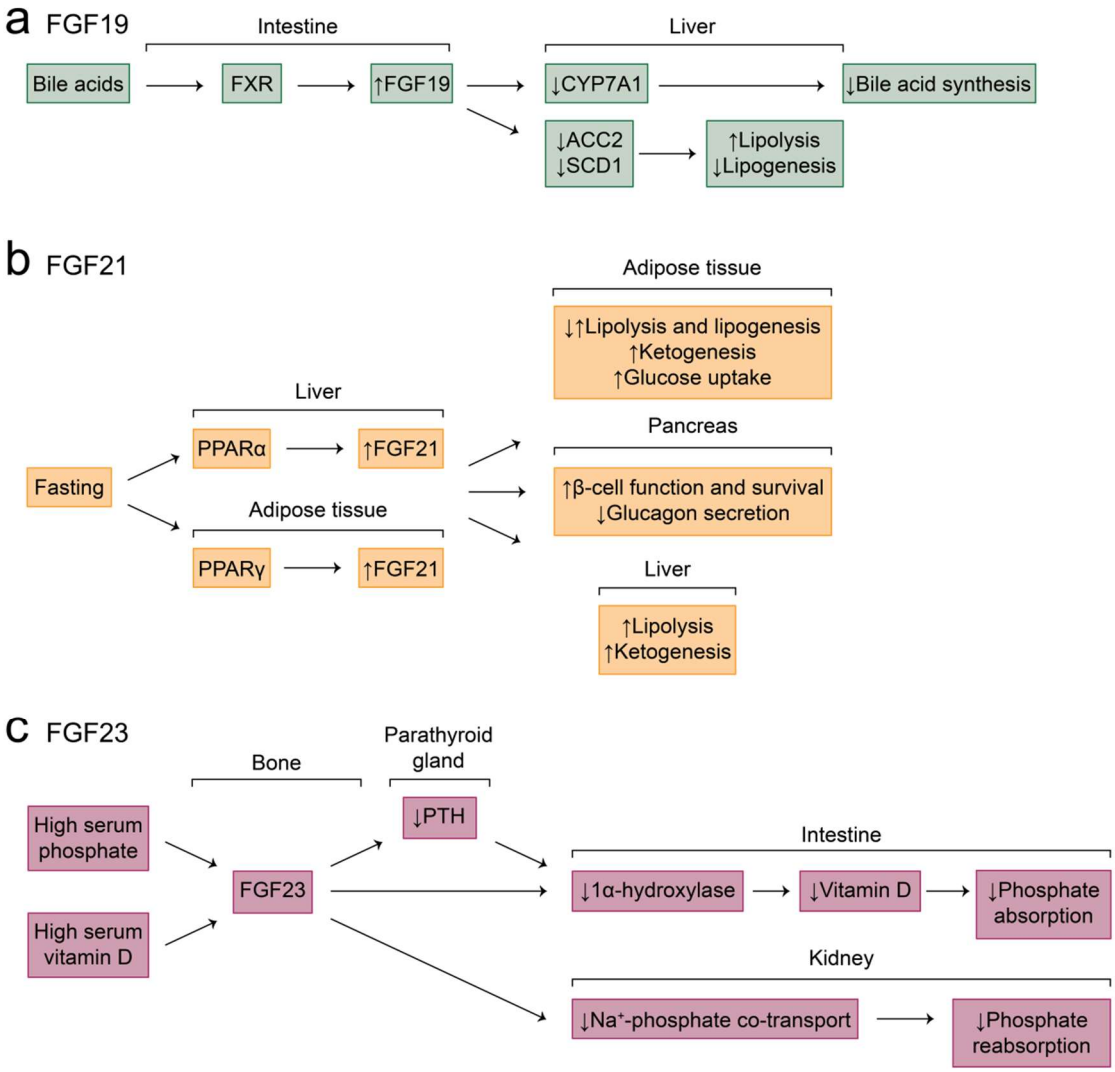


Figure 4.3 Summary of predominant sites of endocrine FGF expression and metabolic action in response to stimuli. Physiological activities of **(a)** FGF19, **(b)** FGF21 and **(c)** FGF23 acting through their respective FGF-FGFR-Klotho ternary complexes (predominantly FGF19-FGFR4-βKlotho, FGF21-FGFR1c-βKlotho and FGF23-FGFR1c-αKlotho). The summary is of literature including aspects of signalling not discussed in detail in the text. FXR – farnesoid X receptor, CYP7A1 – cholesterol 7 α-hydroxylase 1, ACC2 – acetyl-Co-A carboxylase 2, SCD1 – stearoyl-CoA desaturase 1, PPARα/γ – peroxisome proliferator-activated receptor, PTH – parathyroid hormone. Figure reproduced from (Beenken and Mohammadi, 2009).

4.1.8 The ternary complexes and intra-complex protein-protein interactions

The ternary endocrine FGFR signalling complexes comprise specific combinations of FGF, FGFR and Klotho proteins. Endocrine FGFs themselves exhibit lower affinities for FGFRs and HS compared to paracrine FGFs (Harmer et al., 2004, Goetz et al., 2007, Asada et al., 2009, Goetz et al., 2012b), and require Klotho co-receptors for high affinity binding to and signalling via FGFRs (Kurosu et al., 2006, Urakawa et al., 2006, Ogawa et al., 2007, Kurosu et al., 2007, Kharitonenkov et al., 2008, Suzuki et al., 2008). These studies and others have slowly characterised physiologically relevant ternary signalling complexes over a decade of biochemical and cellular research. In general, it was found that αKlotho and βKlotho form mutually exclusive complexes exclusively with their cognate combinations of FGFRs and FGFs, namely FGF23 for αKlotho and FGF19 and FGF21 for βKlotho. Additionally, endocrine FGFs bind with lower affinities to FGFRs or Klotho co-receptors on their own, and bind more effectively to FGFR-Klotho binary complexes. The specific ternary complexes formed regulate distinct physiological outcomes, in a tissue specific manner. They will therefore be presented separately, with a focus on binary and ternary complexes involved in the sub-axes of βKlotho endocrine signalling which is the experimental focus of this work.

4.1.9 FGF23-αKlotho ternary complexes

Direct binding of FGF23 to each of the FGFR isoforms was shown to be weak, with a preference for signalling via the canonical c compared to the alternative b FGFR splice isoforms, and with some dependence on heparin (Yu et al., 2005, Zhang et al., 2006b). The binary interactions between FGF23 and αKlotho, and αKlotho and various FGFRs, and the possible ternary complexes were first demonstrated in HEK 293 cells (Kurosu et al., 2006). This study demonstrated direct binding of αKlotho to various FGFRs, again highlighting a preference for the c compared to the b splice isoforms. FGF23 preferentially bound to binary complexes of αKlotho with either FGFR1c, 3c or 4 (but not 2c) but did not bind to any individual receptor on its own. Additionally, FGF23 signalling could only be activated in the presence of αKlotho, highlighting the Klotho-dependence of FGF23 in receptor binding and signal transduction. Another study similarly demonstrated that αKlotho directly binds FGF23 and the FGFR1c isoform in a HS-independent manner (Urakawa et al., 2006). This study also showed that αKlotho binds to FGFR1c, converting it into a specific HS-independent receptor to which FGF23 can bind and can signal through in a αKlotho-dependent fashion. However, although receptor interaction with FGF23 was strongly αKlotho-dependent and HS-independent, it was suggested that HS might enhance binding and signalling. Two more studies demonstrated that FGF23 could directly bind αKlotho-FGFR1c, but not βKlotho-FGFR1c, to transduce a signal (Suzuki et al., 2008, Kharitonenkov et al., 2008). One study which

demonstrated direct binding of FGF23 to αKlotho (but not βKlotho) *in vitro* also provided evidence that the C-terminal tail residues of FGF23 might be sufficient for binding to αKlotho (Wu et al., 2008). Another study showed that wildtype, but not the N-terminal 'core' of FGF23 without the C-terminal tail, could activate signalling, suggesting the C-terminal tail was necessary and sufficient for signalling (Goetz et al., 2007). Direct *in vitro* competition experiments comprehensively demonstrated that the C-terminal tail of FGF23 does indeed mediate the high affinity interaction of FGF23 with the full length ectodomain complex of αKlotho-FGFR1c, but not with FGFR1c or αKlotho alone, and this occurs in the absence of HS (Goetz et al., 2010). This study also showed that the C-terminal tail of FGF23 can compete with the full length protein for binding to the αKlotho-FGFR1c binary complex and can therefore inhibit signalling. Finally, the binding affinities for the binary interactions between αKlotho and each FGFR isoform were measured *in vitro* by surface plasmon resonance (SPR) (Table 4.2a), confirming that αKlotho binds with highest affinity to FGFR1c, 3c and 4, but not 2c or any of the b splice isoforms (Goetz et al., 2012a). The binding affinities for the interactions between FGF23 and FGFR1c or FGF23 and the αKlotho-FGFR1c binary complex were also measured in comparable experiments using SPR in a second study (Table 4.2b) (Goetz et al., 2012b).

Table 4.2 Equilibrium dissociation constants (K_D, nM) for **(a)** FGFR-Klotho interactions and **(b)** for known FGF-FGFR and FGF-FGFR-Klotho interactions with endocrine FGFs from comparative studies (Goetz et al., 2010, Goetz et al., 2012a, Goetz et al., 2012b).

(a)

	FGFR1c	**FGFR2c**	**FGFR3c**	**FGFR4**
αKlotho	72	–	82	123
βKlotho	124	170	–	84

(b)

	FGF2	**FGF19**	**FGF21**	**FGF23**
FGFR1c	58	1910	2520	648
αKlotho-FGFR1c	–	–	–	27.4

Resolution of the crystal structure of the N-terminal 'core' of FGF23 demonstrated that a large HS-binding region present in paracrine FGFs is missing in FGF23, explaining its reduced affinity for HS and therefore its ability to exert its physiological actions away from its tissue of production by being secreted into the circulation to function as a hormone (Goetz et al., 2007). Mutagenesis experiments performed in this study and a later study (Goetz et al., 2012b) confirmed FGF23 residues involved in HS binding as

the mutations abrogated HS binding of FGF23 further (Figure 4.4c). These mutations to FGF23, which rendered it completely unable to bind HS, did not affect its signalling activity. The latter study also showed that it was possible to engineer FGF2, a paracrine FGF, to act in an endocrine FGF23-like manner by mutating the HS binding residues of FGF2 to abrogate its HS-binding affinity, and adding to it the C-terminal tail of FGF23 enabled the chimeric FGF2 to bind to the αKlotho-FGFR1c complex. So the role of the FGF23 C-terminal tail as the binding site for αKlotho-FGFR1c is emphasised and there is confirmation that this endocrine signalling complex is HS-independent. These studies also confirm that αKlotho (and not βKlotho) is required for FGF23 signalling, and that signalling occurs predominantly via FGFR1c.

4.1.10 FGF19-βKlotho and FGF21-βKlotho ternary complexes

The first reports demonstrating the involvement of βKlotho in FGFR signalling showed that the activities of FGF19 and FGF21 were dependent on the βKlotho co-receptor (Lin et al., 2007, Ogawa et al., 2007). The first study suggested that βKlotho-FGFR4 was the primary receptor for FGF19. Similar to experiments performed for FGF23, the latter study demonstrated binding of βKlotho to FGFR1c and 4, and weakly to FGFR2c. It also showed that FGF21 bound and signalled preferentially through βKlotho-FGFR binary complexes compared to direct signalling through just FGFRs, showing for the first time the βKlotho-dependence of FGF21 signalling. A very similar study by the same group showed that FGF19 also depends on βKlotho, and not αKlotho, for signalling through FGFRs, once again emphasising the importance of FGFR1c and 4, and to a lesser extent FGFR2c and 3c (Kurosu et al., 2007). This study began to address the tissue specificity of signalling through βKlotho-FGFR complexes, showing that both FGF19 and FGF21 can signal in 3T3-L1 adipocytes, which express FGFR1c, but not the other FGFR isoforms. They also showed that only FGF19 could transduce a signal in hepatocytes, which express predominantly FGFR4, although FGF21 was capable of binding to βKlotho-FGFR complexes in these cells. This pattern of signalling in adipocytes and hepatocytes was confirmed *in vitro* and *in vivo* by another study (Adams et al., 2012). The βKlotho-dependence of FGF19 (Wu et al., 2007) and FGF21 (Kharitonenkov et al., 2008) signal transduction respectively were further emphasised by two further studies. One final study demonstrated that both FGF19 and FGF21 could directly bind and signal via the βKlotho-FGFR1c or βKlotho-FGFR3c complexes, and again demonstrated that signalling could not occur through any of the FGFR b splice isoforms (Suzuki et al., 2008). The proposed βKlotho dependent specificity of FGF21 for FGFR1c and dual specificity of FGF19 for FGFR1c and FGFR4 was further evaluated with dose response assays (Gupte et al., 2011) and using dose dependent competition assays (Yang et al., 2012). Tissue- and FGFR-specificity of signalling was further shown using FGFR1c

knockout mouse tissues, where FGF21 signalling was abrogated in adipose tissues (which predominantly express FGFR1c) but FGF19 signalling in hepatocytes (which predominantly express FGFR4) was maintained (Foltz et al., 2012). Direct binding affinities for βKlotho binding to FGFR isoforms were later determined by SPR and confirmed the importance of FGFR1c, 2c and 4, but not 3c or any of the b splice isoforms, in binding to βKlotho (Yie et al., 2009, Goetz et al., 2012a). These studies also observed direct binding of βKlotho respectively to FGF19 and FGF21 (in the absence of FGFRs), a phenomenon which does not occur for FGF23 and αKlotho. The importance of the C-terminal tails of FGF19 and FGF21 in directly binding to βKlotho, but not αKlotho, was first demonstrated by direct binding experiments similar to those performed for FGF23-αKlotho (Wu et al., 2008). Further studies confirmed this, as well as the importance of the C-terminal tails for signalling through βKlotho-FGFR complexes *in vitro* and in cells (Yie et al., 2009, Micanovic et al., 2009, Goetz et al., 2012a). The C-terminal tails of FGF19 and FGF21 were demonstrated to be the βKlotho binding site for these FGFs in similar experiments to those performed for FGF23 by engineering a chimeric FGF, consisting of paracrine FGF2 fused with the C-terminal tail of FGF21, to enable the chimeric FGF2 to bind to βKlotho (Goetz et al., 2012b). Lastly, these studies also suggested that the C-terminal tails of FGF19 and FGF21 compete for the same binding site on βKlotho and bind with variable affinities, implicating the importance of specific sequences of the C-terminal tails in determining ternary complex formation. This also suggests that FGF19 and FGF21 signalling might compete *in vivo* in cells in which other common components of their cognate signalling complexes are expressed, for example in adipocytes.

The role of HS in complexes particularly involving FGF19 has provided substantial controversy. Early work suggested that the FGFR4-HS binary complex might be the unique, HS-dependent and Klotho-independent receptor for FGF19 (Xie et al., 1999). Although a second report suggested that despite FGF19 possibly uniquely binding FGFR4, it was unlikely to bind HS based on the analysis of features of the first crystal structure of FGF19, discussed further below (Harmer et al., 2004). Several reports have shown that although FGF19 signalling is indeed HS-independent, HS might enhance the signal (Zhang et al., 2006b, Lin et al., 2007, Wu et al., 2008, Kharitonenkov et al., 2008) and some reports have also suggested this for signalling via FGF21 and/or FGF23 (Zhang et al., 2006b, Suzuki et al., 2008, Wu et al., 2008). However, some of these studies were performed in cells in which the FGFRs (and some FGFs) are ubiquitously expressed and have used different readouts of activity. Therefore, it could be that HS is enhancing a non-endocrine FGF signalling pathway in these experiments, since paracrine and endocrine signalling can and do occur in parallel *in vivo*. Direct binding of

FGF19, FGF21 and FGF23 to HS *in vitro* has clearly demonstrated the low binding affinity of these FGFs for HS (Goetz et al., 2007, Asada et al., 2009). It is likely that the incorporation of HS in the endocrine signalling complexes could be context dependent, and so far its involvement in endocrine signalling has been reported most frequently for the FGF19-FGFR4-βKlotho complex. Direct binding and activation experiments combined with structural studies with purified ternary complexes are needed to clarify the possible role of HS in enhancing the stability of endocrine signalling complexes.

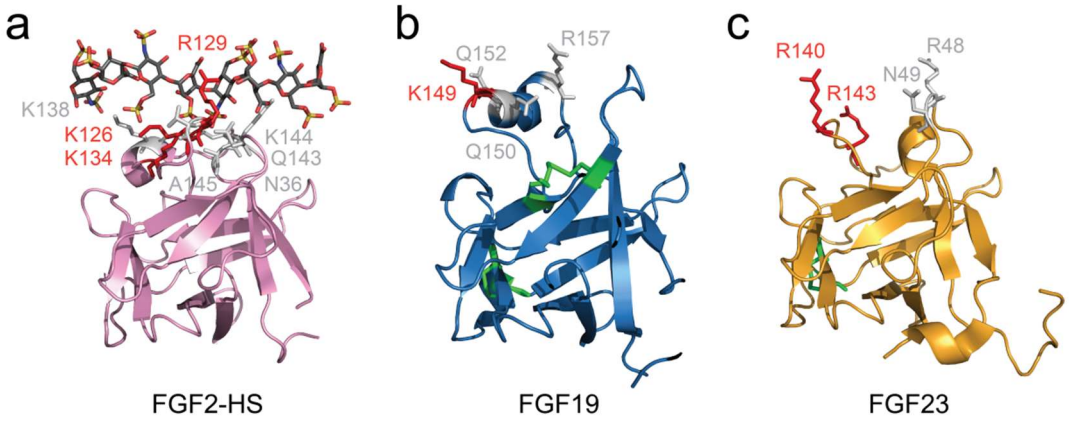


Figure 4.4 Structural basis for the reduced HS binding affinity of endocrine FGFs compared with canonical paracrine FGFs. Comparison of HS binding sites on the crystal structures of human endocrine FGFs (FGF19 and FGF23) with a canonical paracrine FGF (FGF2). **(a)** Crystal structure of paracrine FGF2 (pink cartoon) bound to a heparin decasaccharide (shown as sticks and coloured grey and by atomic element with nitrogen atoms in navy blue, oxygen in red and sulphur in yellow) [PDB: 1FQ9] (Schlessinger et al., 2000). The side chains of HS binding residues are shown as sticks and coloured white and the side chains that form most of the hydrogen bonds with the heparin decasaccharide are coloured red. Crystal structures of endocrine **(b)** FGF19 (blue cartoon) [PDB: 2P23] (Goetz et al., 2007) and **(c)** FGF23 (orange cartoon) [PDB: 2P39] (Goetz et al., 2007) shown in the same orientation as FGF2 in **a**. Side chains of residues that provide residual HS binding affinity are shown as sticks and coloured white, while side chains that abrogate residual HS binding affinity when mutated are shown as sticks and coloured red. Disulphide bonded cysteines in FGF19 (C58-C70 and C102-C120) and FGF23 (C95-C113) are shown in sticks and coloured green. Only one of these disulphide linkages is conserved between the two endocrine FGFs.

As mentioned above, the crystal structure of human FGF19 has been solved, and by two independent groups, yielding two very similar structures (Harmer et al., 2004, Goetz et al., 2007). The structures reveal two disulphide bonds (C58-C70 and C102-C120) in FGF19, compared to none present in all solved structures of paracrine FGFs (with the exception of FGF8, see Figure 1.6b) and only one disulphide bond (C95-C113) in FGF23 (Figure 4.4c). Additionally, mutagenesis experiments similar to those performed for FGF23 (Goetz et al., 2012b) confirmed the residues conferring residual HS binding ability in FGF19 compared to paracrine FGFs (Figure 4.4a-b). The implications of the stabilising disulphide bonds coupled with the reduced HS binding affinities of FGF19 family members have led to the proposal that these two features of endocrine FGFs allow them to diffuse away from their sites of production into the circulation to participate in longer range signalling as hormones.

4.1.11 γKlotho might also form endocrine signalling complexes

Very little is known about the roles of γKlotho in FGFR signalling. A single study has shown that γKlotho might interact specifically with FGFR1b, 1c, 2c, 3c and 4 (Fon Tacer et al., 2010). This study also showed that γKlotho could transduce a signal in the presence of FGF19, but not FGF21 or FGF23, and transduction of this signal was γKlotho-dependent. More evidence will be needed to elucidate the functional roles of this member of the Klotho family in FGF signalling.

4.1.12 Specific interactions between FGFs, FGFRs and Klotho co-receptors

The FGF23 C-terminal tail was shown to confer high affinity binding to αKlotho-FGFR binary complexes, and the C-terminal tails of FGF19 and FGF21 were shown to directly bind with high affinity to βKlotho and βKlotho-FGFR binary complexes, discussed in detail in sections 4.1.9 and 4.1.10. However, so far there is no information about specific FGF or FGFR binding sites on αKlotho and βKlotho. Additionally, limited molecular detail has been elucidated regarding the Klotho-FGFR and endocrine FGF-FGFR interfaces and the specific sequences involved in their binding interactions. The binding affinities for the interactions between FGFR isoforms and Klotho co-receptors have been measured *in vitro* using the ectodomains of each receptor (Table 4.2a). These support the possibility of heterodimerisation of these receptors, and thus binary complex formation, prior to the addition of ligand to form ternary complexes.

The first structural information to date regarding specific binding interfaces comes from a few main studies (Gupte et al., 2011, Wu et al., 2011, Goetz et al., 2012a). The first study created chimeric FGFR1c and FGFR4 receptors by exchanging either their D2 or D3 domains, demonstrating that the D3 domain provided the specific determination of

binding and activation by FGF21 in the presence of βKlotho, while FGF19 activation was less affected by swapping of the D3 domains but was slightly affected when the D2 domain of FGFR4 was removed in exchange for the FGFR1c D2 domain. They went further to propose a potential six amino acid sequence in the βC'-βE loop of the D3 domain of FGFR1c as a potential FGF21 interaction site (residues 318-323, 'TTDKEM'). The second study implicated the N-terminus of FGF19 in binding to FGFR4. The third study also mapped a partial Klotho-binding site to residues of the hydrophobic 'groove' in the D3 domain of FGFR1c (Figure 4.5). This was achieved by demonstrating that αKlotho, βKlotho and the paracrine FGF8b could all bind to an FGFR1c ectodomain protein comprising the D2-D3 domains, but specific site-directed mutagenesis of hydrophobic residues in the D3 domain 'groove' (conserved across FGFR1-3 c splice isoforms and FGFR4) diminished binding of all three proteins to the FGFR1c protein. Furthermore, FGF8b signalling was inhibited in cells in the presence of βKlotho, suggesting overlap between the FGF8b and Klotho binding sites on FGFR. This has important tissue specific implications is suggesting that paracrine and endocrine signalling must complete with each other in tissues in which multiple FGFs, FGFRs and Klotho proteins are present. Endocrine signalling might 'win' in tissues in which the correct combinations of FGFRs and Klotho proteins are expressed, or perhaps highly regulated thresholding of levels of expression of the Klotho proteins might play a role in modulating signalling. Additionally, a recent study monitoring the effects of excess levels of FGF19 and FGF21 in mice suggested the two hormones might compete for binding to the βKlotho-FGFR4 complex to exert opposite effects on inhibiting and stimulating bile acid synthesis respectively (Zhang et al., 2017a). This study proposed that FGF21 achieves this through direct inhibition of FGF19 by binding to, but not signalling through, the βKlotho-FGFR4 binary complex. This suggests a role for crosstalk between endocrine FGFs in regulating metabolism, similar to that observed for FGF8b and βKlotho in modulating crosstalk between paracrine and endocrine signalling (Goetz et al., 2012a). These studies suggest an additional method of functional regulation of endocrine signalling through competition between FGFs for Klotho-FGFR complexes and between FGFs and Klotho co-receptors for binding to FGFRs, ultimately mediating the output effect on metabolism. However, spatiotemporal expression of different factors cannot be ruled out as a possible contributor to the outcome of signalling via these factors *in vivo*, and all these hypotheses remain to be investigated further.

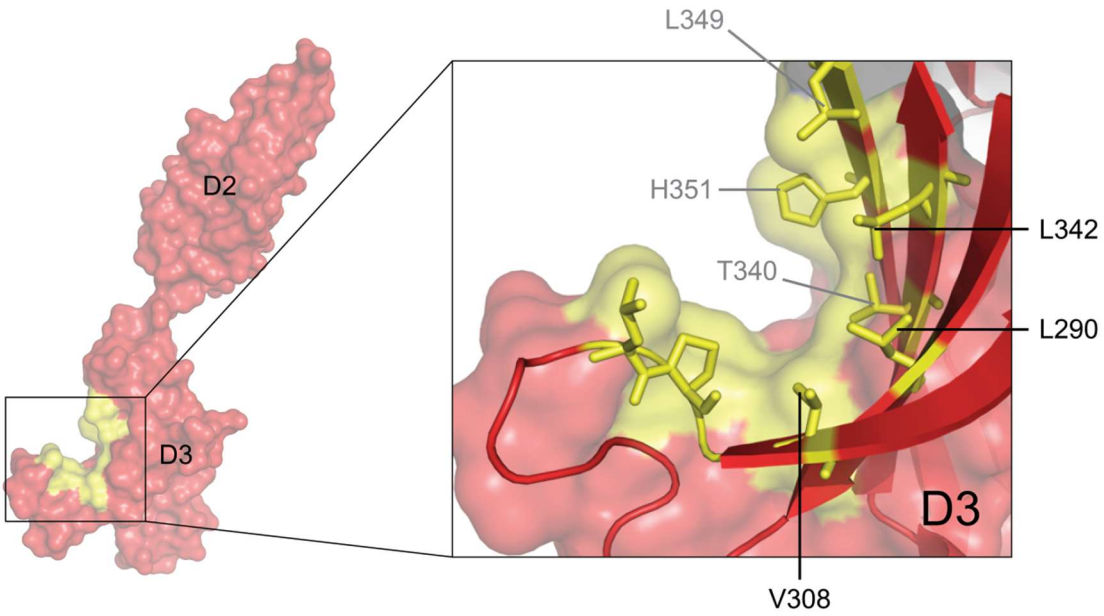


Figure 4.5 Location of a potential Klotho binding region in a hydrophobic 'groove' in the D3 domain of FGFR1c. Crystal structure of the ligand binding region of FGFR1c (D2-D3) shown as a surface and coloured red with the hydrophobic 'groove' coloured yellow (left panel) [PDB: 1CVS] (Plotnikov et al., 1999). A close-up view of the D3 domain is shown as a cartoon against a surface and coloured red (right panel). The residues that make up the hydrophobic 'groove' in FGFR1c are shown as sticks against a surface and coloured yellow. Mutations at three residues shown to most strongly abrogate binding to αKlotho, βKlotho and FGF8b (Goetz et al., 2012a) are labelled in black. Mutations at other residues also shown to perturb binding to αKlotho and βKlotho to a lesser degree are labelled in grey.

A therapeutic antibody has been found to compete with FGF21 for binding to βKlotho, and the discontinuous region on the KL1 domain of βKlotho implicated in binding encompasses residues 1-80 and 303-522, proposing that the FGF21 binding region on βKlotho might be (at least in part) on the KL1 domain (Foltz et al., 2012). It has also been proposed that the intracellular C-termini of βKlotho and FGFR1c might provide part of their binding interfaces (Ming et al., 2012), however this has not been directly demonstrated experimentally and is unlikely due to the very short intracellular domains of the Klotho proteins (Figure 4.2 and Table 4.1). Another study has proposed the D1 domain of FGFR1c as an inherent inhibitor of formation of the FGF21-FGFR1c-βKlotho complex, while implicating the N-terminus of FGF21 in binding to the FGFR1c D2-D3 region (Yie et al., 2012). Further work is needed to define the true and specific interfaces between all receptor components of the ternary complexes.

4.1.13 Stoichiometry of components of the ternary complexes

So far, almost no experimental evidence exists defining the composition of the ternary signalling complexes with respect to the stoichiometry of proteins within the complexes. Three main working models for the functional ternary complexes were proposed, namely a 2:2:2, 1:2:1 and a 2:2:1 model of FGF-FGFR-Klotho components (Goetz and Mohammadi, 2013, Kilkenny and Rocheleau, 2016). Various studies have proposed similar models and some evidence has emerged supporting either the 2:2:2 complex (Yie et al., 2012, Comps-Agrar et al., 2015) or the 2:2:1 complex (Ming et al., 2012). It has also been suggested that the galectin lattice might play a role in mediating Klotho protein availability at the cell surface, particularly for βKlotho which has the most glycosylation sites of all Klotho family members (Figure 4.2 and Table 4.1), thus modulating complex formation as reviewed in (Kilkenny and Rocheleau, 2016). Additionally, the fact that Klotho proteins can bind to FGFRs in the absence of FGF ligands suggests that pre-formed Klotho-FGFR heterocomplexes might exist (if only transiently) prior to ligand binding, but the possible stoichiometries of such pre-formed heterocomplexes has only been partially and indirectly evaluated so far (Figure 4.6a). It is likely that all proposed stoichiometric models might exist in different contexts. Structural information at the molecular level is therefore much needed to resolve the true nature of the stoichiometry of components of the ternary complexes.

4.1.14 Summary of endocrine signalling complexes and interactions

Several studies have begun to elucidate the compositions and functional significance of Klotho-dependent endocrine signalling complexes. These studies have shown that signalling via FGF19 and FGF21 is βKlotho-dependent, in combination with different possible FGFR isoforms, while FGF23 signalling is αKlotho-dependent and occurs mostly through FGFR1c. The importance of FGFR1c and 4 has been rigorously confirmed, while the roles of FGFR2c and 3c in endocrine signalling remain to be deciphered and are likely context dependent. The predominant modes of action of the endocrine signalling complexes believed to be functionally important to date are summarised in Figure 4.3. Working models for proposed stoichiometries of the ternary complexes are summarised in Figure 4.6b In general, the tissue specific expression of Klotho family members and FGFR isoforms *in vivo* dictates the tissue selectivity and context dependence of signalling and metabolic activities of endocrine FGFs.

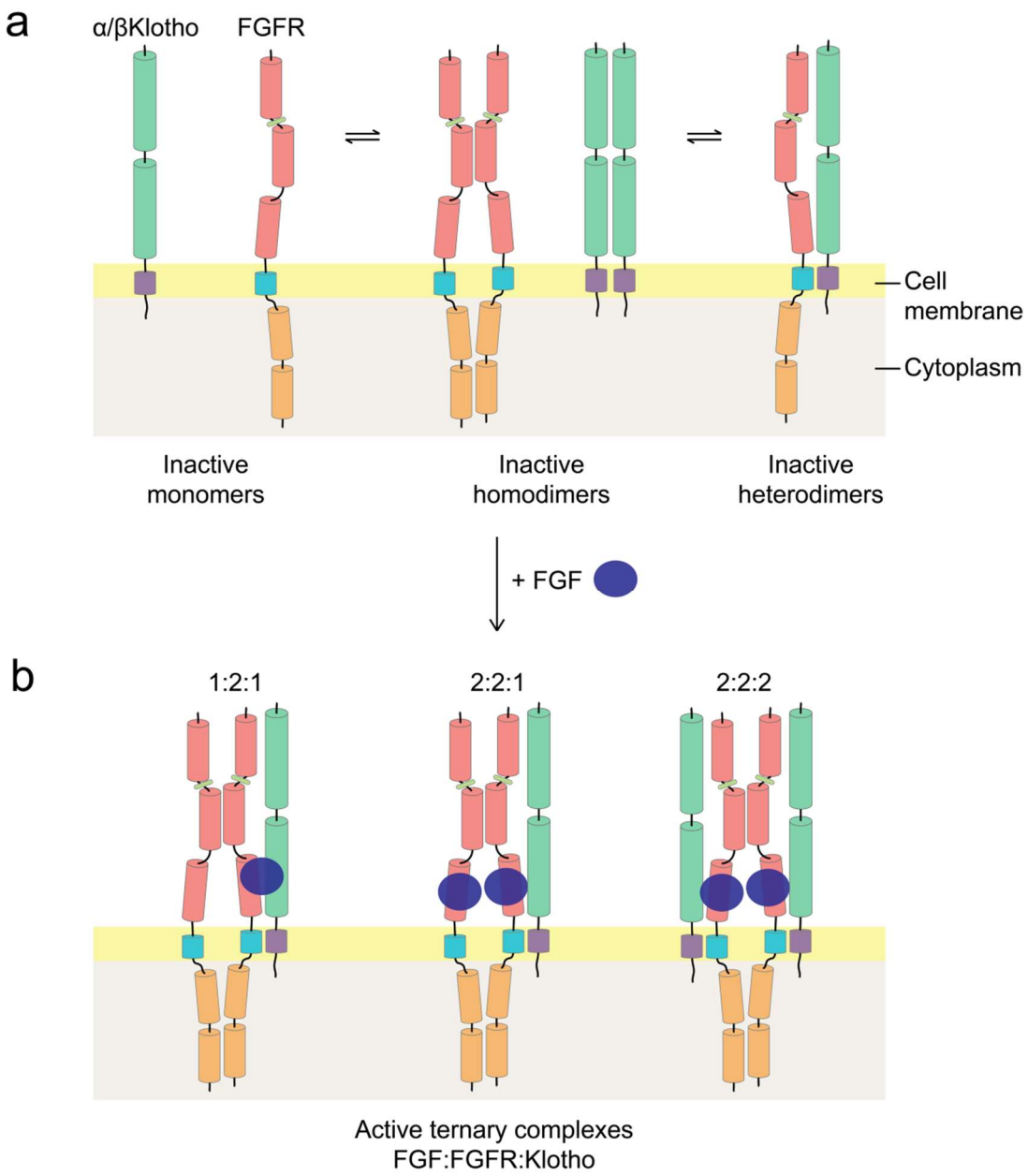


Figure 4.6 Current working models for receptor stoichiometries of endocrine ternary signalling complexes, before and after FGF ligand binding. (a) Possible receptor stoichiometries at the cell surface for FGFRs and α/βKlotho prior to FGF conjugation. Both FGFRs and Klotho co-receptors might exist in equilibrium between monomeric, homodimeric and heterodimeric inactive states. Dimeric states might represent 'primed' conformations for binding to FGFs. **(b)** Stoichiometries of proposed functional endocrine ternary signalling complexes, demonstrating the components of 1:2:1, 2:2:1 and 2:2:2 FGF:FGFR:Klotho complex stoichiometries. The role of heparan sulphate in these complexes is unknown and is not shown.

4.1.15 Age related pathologies predominantly involving αKlotho

Overexpression of the αKlotho ectodomain has been shown to suppress ageing and extend the lifespan of mice by specifically inhibiting insulin and insulin growth factor 1 (IGF1) signalling, with no effects on EGFR or PDGFR (Kurosu et al., 2005). αKlotho overexpression in breast cancer cell lines first revealed its potential tumour suppressive properties (Wolf et al., 2008). It has since been implicated as a tumour suppressor in a range of cancers (Zhou and Wang, 2015). αKlotho is often depleted in early stage cancers due to epigenetic silencing of the *Klotho* gene by promoter hypermethylation as first shown in cervical carcinoma (Lee et al., 2010) and subsequently in a range of other cancers, reviewed in (Mencke et al., 2017). The protective and tumour suppressive roles of αKlotho in cancer are increasingly being evaluated (Xie et al., 2013) alongside its possible benefits as a biomarker for several diseases (Zhou and Wang, 2015). Similarly, βKlotho has independently been suggested to act as a tumour suppressor through FGF19-induced apoptosis of hepatoma cells (Luo et al., 2010), although this is in contrary to a much larger pool of evidence implicating FGF19 as a mitogen in hepatocellular carcinomas (discussed in section 4.1.19). The consequences of αKlotho deficiency have been implicated in a range of pathologies including acute kidney injury (Hu and Moe, 2012), chronic kidney diseases (Hu et al., 2013, Hu et al., 2012) and cancer and fibrosis (Mencke et al., 2017). FGF23 has also been found to be associated with chronic kidney disease and cardiovascular diseases (Kovesdy and Quarles, 2016), and all this evidence alludes to the FGF23-FGFR1c-αKlotho complex as a therapeutic target for several important age-related pathologies.

4.1.16 Germline diseases caused by mutated FGF23

Congenital gain-of-function mutations in *FGF23* lead to the development of autosomal dominant hypophosphatemic rickets through stabilisation of the active form of FGF23 (ADHR Consortium, 2000, White et al., 2001) while loss-of-function mutations lead to hyperphosphatemic familial tumoural calcinosis (Benet-Pages et al., 2005, Chefetz et al., 2005, Garringer et al., 2008). FGF23 also causes tumour-induced osteomalacia in humans (Shimada et al., 2001) and has once been implicated to possibly stimulate the progression of prostate cancer (Feng et al., 2015).

4.1.17 Autocrine FGF19 signalling is an oncogenic driver in liver cancers

It was found that FGF19, but not FGF21, could induce the proliferation of hepatocytes in an FGFR4-dependent manner (Wu et al., 2010, Wu et al., 2011). The importance of FGF19 in liver regeneration has also been confirmed in mice (Uriarte et al., 2013, Kong et al., 2014). The chromosomal locus harbouring the *FGF19* gene has more recently been found amplified in hepatocellular carcinoma (and relevant liver cancer cell lines) by

several studies (Sawey et al., 2011, Guagnano et al., 2012, Kaibori et al., 2016). FGF19 has also been found to act as a paracrine or autocrine stimulator of proliferation in hepatocellular carcinoma (Nicholes et al., 2002, Uriarte et al., 2015), as well as in other tumours such as in prostate cancer (Feng et al., 2013), ovarian cancer (Hu and Cong, 2015) and hepatoblastoma (Elzi et al., 2016). A recent study demonstrated that amplification of *FGF19* is the oncogenic driver for a subset of hepatocellular carcinoma which is dependent on FGF19 signalling via FGFR4 (Sawey et al., 2011). Amplification of *FGF19* has also been shown to confer sensitivity to targeted FGFR inhibitors in specific subsets of liver cancer cell lines, but only when the βKlotho co-receptor and FGFR4 are co-expressed (Guagnano et al., 2012, Futami et al., 2017). Another study showed that chronic exposure to FGF19 for a year in mice can lead to hepatocellular carcinoma formation, but not if the N-terminal residues 23-42 and 50-57 are substituted by the corresponding residues of FGF21 (Ge et al., 2012). Together, this work strongly implicates the FGF19-FGFR4-βKlotho complex as an oncogenic driver and therapeutic target in hepatocellular (and possibly other) carcinomas, particularly for a subset of nearly a third of hepatocellular carcinomas which express FGF19, FGFR4 and βKlotho.

4.1.18 Clinical adverse effects involving endocrine FGFR signalling

Anti-FGFR targeting in the clinic, involving both multi-kinase and FGFR-specific inhibitors, has slowly progressed over recent years, as discussed in Chapter 1. Large advances have been made to develop potent and specific pan-FGFR inhibitors that predominantly target FGFR1-3 (Table 1.4). However, in contrast to multi-kinase inhibitors for which clinical toxicity normally involves off-target adverse effects via co-inhibition of other RTKs (Dieci et al., 2013, Shah et al., 2013b, Dienstmann et al., 2014, Soria et al., 2014), it has been widely observed that the adverse side-effects in response to potent and specific pan-FGFR inhibitors in pre-clinical tumour models are most commonly due to on-target inhibition of endocrine FGFR signalling (Herbert et al., 2014). Hyperphosphatemia-mediated tissue calcification is the most common adverse outcome in pre-clinical *in vivo* models due to inhibition of FGF23 signalling (Brown et al., 2005, Dieci et al., 2013, Wohrle et al., 2013). It is expected that this and other side-effects will present in the clinic in response to FGFR-targeting with the most potent inhibitors, but early clinical data suggests that the adverse effects might be manageable and reversible by careful management (Nogova et al., 2017, Dienstmann et al., 2014). Better understanding of endocrine FGF23 signalling will help to create targeted agents to re-activate signalling via this crucial pathway during inhibition of oncogenic FGFR signalling in the clinic, since adverse side-effects have so far prevented optimal anti-cancer targeting due to premature termination or dose reduction of the targeted therapy in the case of severe adverse effects (di Martino et al., 2016).

4.1.19 Endocrine FGFs as biologic therapeutics and biomarkers

Involvement of endocrine signalling (predominantly through FGF19 and FGF23) in a range of cancers and inherited and metabolic diseases has led to interests in these molecules and their ternary signalling complexes as therapeutic targets. However, potential safety concerns of targeting FGF19-FGFR4 signalling in hepatocellular carcinoma have been emphasised due to possible adverse on-target effects such as on bile acid synthesis (Mellor, 2014). Moreover, the potential anti-diabetic, anti-obesity and life extending effects of FGF21 (which additionally does not exhibit the mitogenic activities seen with FGF19) have arguably fostered the most therapeutic interest in developing this hormone as a biologic therapeutic compared to the other two endocrine FGFs. One therapeutic antibody for X-linked hypophosphatemia (caused by elevated FGF23 levels) is the only endocrine FGF related therapy to date to reach Phase III clinical trials. A plethora of patents has been filed covering a multitude of variants and fusions of endocrine FGFs, therapeutic antibodies and various other agonists and antagonists as context dependent therapeutics. A current selected list of endocrine FGF related therapies at various stages of pre-clinical or clinical development is summarised in Table 4.3. Additionally, the levels of endocrine FGFs have been found to be altered in the blood circulation of patients with various metabolic, cardiovascular and oncogenic diseases, reviewed in (Itoh and Ornitz, 2011, Itoh et al., 2015). Endocrine FGF levels in patients are therefore thought to serve as useful biomarkers for several metabolic, cardiovascular and malignant diseases.

Table 4.3 Selected endocrine FGF related therapeutics currently in pre-clinical or clinical development. mAb – monoclonal antibody, HSA – human serum albumin, PEG – polyethylene glycol, NYU – New York University, WMC – Wenzhou Medical College

Name	Description	Company	Reference
1A6	Anti-FGF19 mAb	Genentech	(Desnoyers et al., 2008)
LD1	Anti-FGFR4 mAb	Genentech	(French et al., 2012)
U3-1784	Anti-FGFR4 mAb	Daiichi Sankyo	Clinical Trial NCT02690350
sFGFR4	Soluble FGFR4 mimic, dominant negative	–	(Ezzat et al., 2001)
BLU9931	FGFR-paralogue specific irreversible anti-FGFR4 small molecule inhibitor	Blueprint Medicines	(Hagel et al., 2015)
NVP-FGF401	Anti-FGFR4 small molecule inhibitor	Novartis	Clinical Trial NCT02325739
BLU-554	Anti-FGFR4 small molecule inhibitor	Blueprint Medicines	Clinical Trial NCT02508467
H3B-6527	Anti-FGFR4 small molecule inhibitor	H3 Biomedicine	Clinical Trial NCT02834780
FGF21	Native FGF21	–	–
LY2405319	Aggregation-resistant, improved stability FGF21 analogue (ΔHPIP, L118C, A134C, S167A)	Eli Lilly	(Kharitonenkov et al., 2013)
FGF21/19	FGF21/19 chimera	NYU	(Goetz et al., 2012a)
$FGF21^{1-171}$-WD22	$FGF21^{1-171}$-WD22 avimer	Amgen	(Smith et al., 2013b)
Fc-FGF21	Aggregation-resistant, improved stability, long acting FGF21 analogue (Fc^{IgG1}, L98R, P171G)	Amgen	(Hecht et al., 2012)
PEG-FGF21	PEGylated FGF21 analogues	Amgen	(Xu et al., 2013)
PEG-FGF21	PEGylated FGF21 analogue	WMC	(Huang et al., 2011)
ARX-618	PEGylated FGF21 analogue	Merck	(Mu et al., 2012)
CVX-343	FGF21 (ΔH1, A129C) covalently linked to IgG1κ mAb CVX200	Pfizer	(Huang et al., 2013a)
FGF2/21	FGF2/21 chimera	NYU	(Goetz et al., 2012a)
mimAb1	Anti-βKlotho / anti-FGFR1c-βKlotho mAb (FGF21 mimetic)	Amgen	(Foltz et al., 2012)
C3201-HSA	HSA-coupled FGFR1c-βKlotho bispecific avimer	Amgen	(Smith et al., 2013a)
bFKB1	Anti-FGFR1c-βKlotho bispecific agonist mAb	Genentech	(Kolumam et al., 2015)
KRN23	Anti-FGF23 neutralising mAb	Ultragenyx	(Yamazaki et al., 2008)
FGF23 C-tail	FGF23 C-terminal tail peptides	NYU	(Goetz et al., 2010)
Klotho-FGF	Klotho-FGF fusion peptides	Novartis	–
sKlotho variants	sKlotho variant polypeptides	Novartis	–
F91-8A07	Artificial peptide agonist to FGFR1c-βKlotho	Takeda	(Sakamoto et al., 2016)

4.2 Aims

Deregulated FGFR signalling is implicated in diverse aspects of tumourigenesis (section 4.1.17 and Chapter 3) and in mediating on-target toxic side effects observed in response to treatment with potent and specific FGFR inhibitors in the clinic (section 4.1.18). However, despite their importance in modulating these effects, the endocrine sub-axes of FGFR signalling have so far been less well-studied than the 'canonical' paracrine modes of FGFR signalling presented in Chapter 1. Although considerable advances have been made in understanding the physiological roles of endocrine FGFR signalling in biology and pathophysiology, less is known about the molecular details in terms of structures, stoichiometry and protein-protein interactions of the ternary endocrine signalling complexes, and there are in fact no atomic structures for any of the Klotho proteins. Together, this provides a strong rationale to further study endocrine FGFR signalling proteins and complexes from a structural, biophysical and biochemical perspective.

This work sought to characterise endocrine FGFR signalling complexes involving βKlotho, its cognate ligands FGF19 and FGF21 and FGFR1c from a structural and biophysical perspective. To this end, bioinformatics analyses, homology modelling and analyses of existing crystal structures and the literature were used to design constructs for protein expression. Methods were then developed and optimised to express and purify protein components of the binary and ternary endocrine signalling complexes, and to assess the 'foldedness' and bioactivity of refolded recombinant proteins. Finally, preliminary biophysical experiments were performed to begin to understand the protein-protein interactions involved in complex formation. This work will inform the rational design of therapies already in development for targeting deregulated endocrine FGFR signalling in various metabolic disorders and cancer.

4.3 Results

4.3.1 Structural bioinformatics analyses of human βKlotho

As discussed in section 4.1, there is a wealth of existing structural and biochemical information available for FGFR1 and little information about some endocrine FGFs. However, there is limited biochemical and no structural information available in the public domain for any of the full length Klotho proteins or their individual KL domains. Therefore to determine potential N- and C-terminal boundaries for βKlotho protein constructs to be recombinantly expressed in this work, there was a requirement for some *in silico* predictive structural bioinformatics analyses prior to construct design. A combination of tools including multiple sequence alignment, secondary structure prediction and homology modelling were utilised for this purpose. This analysis was then compared with constructs available commercially and from the literature.

Sequence alignments of the predicted βKlotho KL1 and KL2 domains, as well as several other sequence alignments of human, mouse and rat αKlotho and βKlotho protein sequences were performed using the online multiple sequence alignment tool Clustal Omega (Sievers et al., 2011). It was found that human and mouse βKlotho share 79% sequence identity, with their individual KL domains sharing over 80% sequence identity. Human αKlotho and βKlotho share just 45% sequence identity, and human βKlotho KL1 and KL2 domains only share 28% sequence identity (30% for the orthologous mouse domains). A Basic Local Alignment Search Tool (BLAST) search (Altschul et al., 1990) combined with assessment of the human KLB UniProtKB entry with accession #Q86Z14 (The UniProt, 2017), and secondary structure predictions including predictions of secretion signal peptide and transmembrane regions using online tools SignalP (Petersen et al., 2011) and TMHMM (Krogh et al., 2001) helped in defining potential domain lengths and boundaries for the individual and combined KL domains. Construct boundaries used previously in commercially available recombinant proteins (R&D Systems, #5889-KB) and in the literature (Goetz et al., 2012a) were also considered. These sequence analyses are summarised in Figure 4.7 and Table 4.4.

Homology models of human βKlotho KL1 and KL2 domains were generated separately using different templates due to the low sequence identity and similarity (28% and 36% respectively) between these two domains. A cyanogenic β-glucosidase (PDB: 1CBG, resolution 2.15 Å) (Barrett et al., 1995) was found to share 40% sequence identity with residues 66-505 of the KL1 domain. A cytosolic β-glucosidase, also called Klotho related protein (KLrP) (PDB: 2E9M, resolution 1.80 Å) (Hayashi et al., 2007) was found to share 38% sequence identity with residues 521-967 of the KL2 domain. These two templates

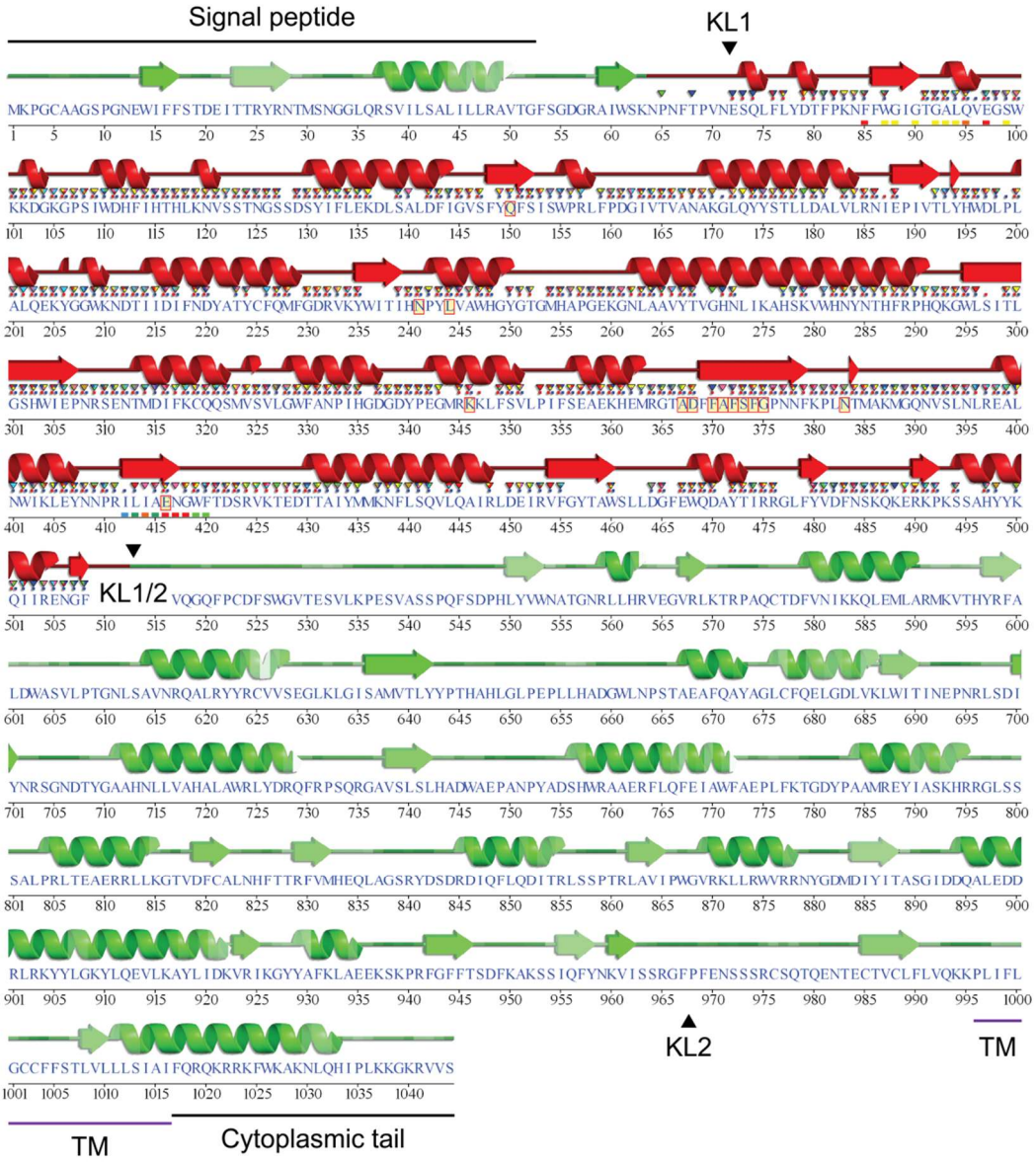


Figure 4.7 Summary of secondary structure prediction for human βKlotho. A compilation of several secondary structure prediction analyses performed in this work for full length human βKlotho (accession #Q86Z14)). Annotations have been displayed on the predicted sequence-structure output from the Sequence Annotated by Structure (SAS) online tool (Milburn et al., 1998). A putative signal peptide and transmembrane region are labelled, as predicted using online tools SignalP (Petersen et al., 2011) and TMHMM (Krogh et al., 2001). The short cytoplasmic domain (see Table 4.1) is also labelled. The predicted KL1 domain is coloured red based on sequence homology to β-glucosidases. Putative boundaries of KL1 and KL2 domains predicted by a combination of methods (Table 4.4) are marked with black triangles and labelled. Accession numbers for entries from the UniProtKB database for other αKlotho and βKlotho sequences used for multiple sequence alignments include human αKlotho (#Q9UEF7), mouse αKlotho (#O35082), rat αKlotho (#Q9Z2Y9) and mouse βKlotho (#Q99N32). For full details refer to text and see Table 4.4.

Table 4.4 Extracellular βKlotho domain boundary annotations curated using *in silico* bioinformatics tools BLASTP (Altschul et al., 1990), UniProtKB (The UniProt, 2017) using accession #Q86Z14 (human KLB), Clustal Omega (Sievers et al., 2011), SAS (Milburn et al., 1998), SignalP (Petersen et al., 2011) and TMHMM (Krogh et al., 2001) or from expression constructs supplied from a commercial source (R&D Systems) or used in the literature (Goetz et al., 2012a).

Source	Orthologue	Predicted or Expressed?	KL1 domain	KL2 domain
BLASTP + UniProtKB	Human	Predicted	77-508	517-967
Clustal Omega (multiple sequence alignments)	Human, mouse, rat	Predicted	73-513	512-972
SAS (secondary structure prediction)	Human	Predicted	73-508	513-962
SignalP + TMHMM	Human	Predicted	53-996	
Commercial (R&D Systems, #5889-KB)	Human	Expressed (mouse myeloma cell line)	53-997	
Literature	Mouse	Expressed (HEK 293 cell line)	53-995	

generated the best-scoring homology models of KL1 and KL2 domains respectively (see section 2.19 for more details). The model structures are overlaid and compared in Figure 4.8a. The KL1 and KL2 domain models appear structurally similar, despite their sequence diversity. They adopt a TIM-barrel fold predominantly consisting of eight parallel β-strands surrounded by eight stabilising α-helices forming a central core (Figure 4.8c), similar to the structures of cyanogenic β-glucosidase and KLrP proteins from which the structures were modelled. The α-helices and β-strands alternate along the amino acid sequence and are occasionally interrupted by other small α-helix or β-strand insertions (Figure 4.8c). From these models, the spatial distribution of the two (KL1) and five (KL2) cysteine residues in each KL domain relative to each other (Figure 4.8b) suggests that intra-domain disulphide bonding is unlikely to occur. It is difficult to predict from these models of individual KL domains whether any of the cysteines might be involved in inter-domain or inter-molecular disulphide bonding. Only one cysteine is conserved between the KL1 and KL2 domains (C225 in KL1 and C677 in KL2) both in sequence alignment and in the structural model, but it is buried in the core structure and is not solvent exposed (Figure 4.8b). It is therefore unclear whether it might have any functional role such as in inter-molecular dimer formation.

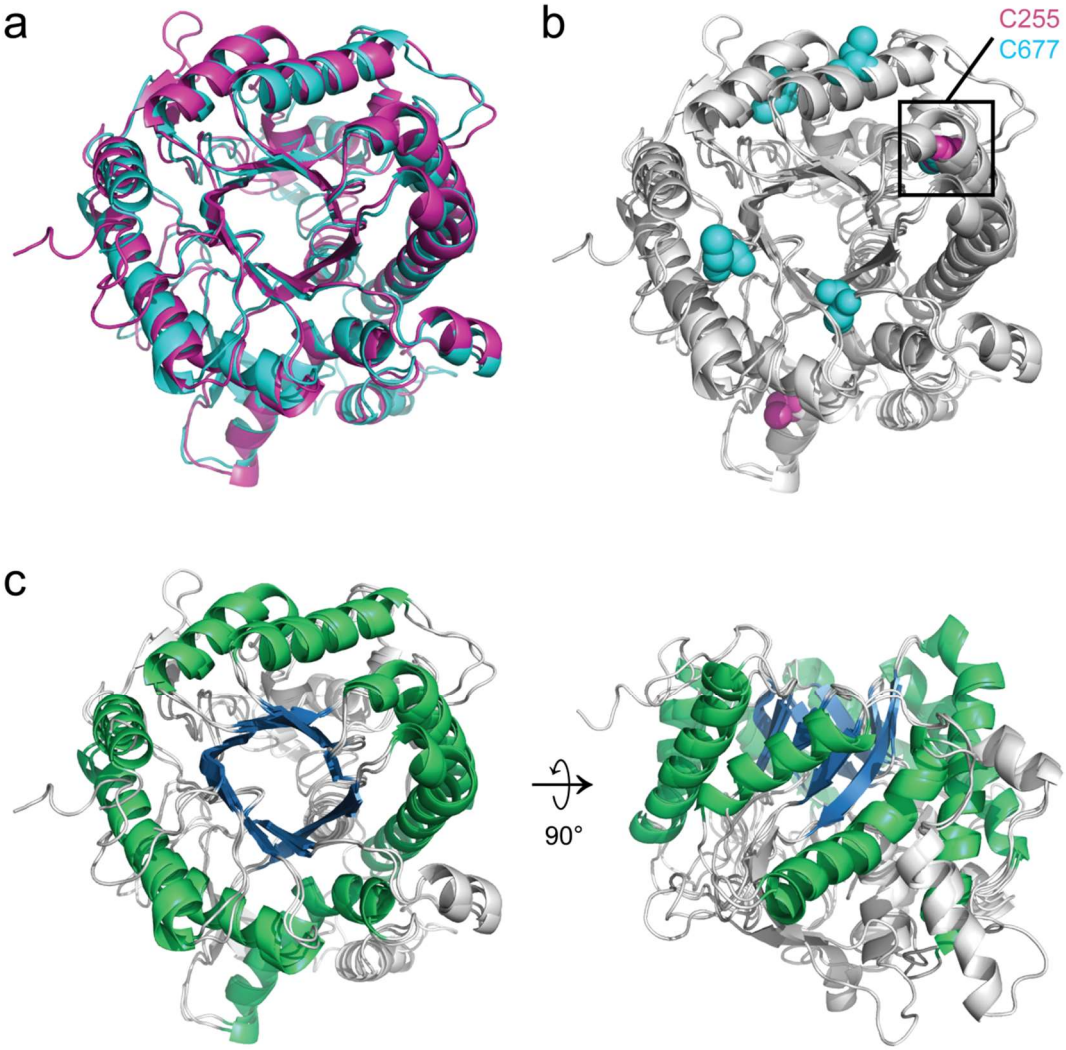


Figure 4.8 Homology models of human βKlotho KL1 and KL2 domains. (a) Model of human βKlotho KL1 (pink) overlaid with a model of human βKlotho KL2 (cyan) generated using the SWISS-MODEL web-based protein structure modelling workspace (Arnold et al., 2006). Two different templates were used for each KL domain, each with the highest sequence identity to each individual KL domain, selected from searches performed against the template library within the SWISS-MODEL workspace. Despite low (28%) sequence identity between the two KL domains, the homology models are very similar with an RMSD of 1.153. For full experimental details refer to text and see methods section 2.19. **(b)** Overlay of the same KL1 and KL2 models with all cysteine resides depicted as spheres for KL1 (pink) and KL2 (cyan), demonstrating that only one cysteine is conserved between the KL1 and KL2 domains (boxed and labelled) whose function is unknown. All cysteines within each KL domain are further apart than 2.05 Å (the length of a disulphide bond) so intra-domain disulphide bonding is unlikely to occur. However, it is difficult to infer putative intermolecular (between βKlotho monomers) or KL1-KL2 inter-domain (within the same βKlotho monomer) disulphide bonds from these models. **(c)** Overlay of the same KL1 and KL2 models highlighting their TIM-barrel structural fold. Eight parallel β-strands (blue) are surrounded by eight stabilising α-helices (green) forming a central core, with several insertions in between. Left and right panels show the same overlay from two different perspectives rotated relative to each other by 90° around the horizontal axis as depicted.

Based on these combined analyses, it seemed most probable that βKlotho KL domains could reasonably be defined by boundaries 73-508 and 513-967 for the KL1 and KL2 domains respectively, separated by a short KL1-KL2 linker. However, the N-terminal KL1 boundary appears to be very close to a region of predicted secondary structure (Figure 4.7). Additionally, the signal peptide is predicted to terminate between residues 52-53, therefore it is likely that the true N-terminus of the mature βKlotho protein exists at residue 53, supported by the constructs used in the literature (Table 4.4). The C-terminus of the extracellular region is predicted to terminate at residue 996. Therefore, based on all this information, the construct boundaries for the βKlotho KL1 and KL2 protein constructs for bacterial expression in this work were defined as 53-514 and 513-972 respectively, and as 53-995 for the entire extracellular KL1-KL2 construct. These incorporate each entire KL domain surrounded by a few extra linker residues and are comparable to constructs produced by commercial sources and used in the existing literature.

4.3.2 Bacterial expression constructs: design and expression testing in *E. coli*

In order to reconstitute FGFR endocrine signalling complexes *in vitro*, purification of the three components of the ternary protein complexes (FGF, FGFR and Klotho) were required. Previous efforts by predominantly one research group have shown that human and mouse FGF19, FGF21 and extracellular FGFR1 proteins can be expressed and refolded from inclusion bodies produced in *E. coli*, for example (Plotnikov et al., 1999, Harmer et al., 2004, Kharitonenkov et al., 2005, Ryu et al., 2006, Wu et al., 2008, Asada et al., 2009, Xu et al., 2009, Yie et al., 2009, Beenken et al., 2012) while other groups have chosen to express the endocrine FGF proteins in diverse cellular systems including in mammalian, insect and yeast cell lines (Xie et al., 1999, Inagaki et al., 2005, Micanovic et al., 2009, Wu et al., 2011, Song et al., 2016). Some attempts have also been made to express and/or purify the ectodomains of the human βKlotho (Yie et al., 2009), mouse αKlotho (Kurosu et al., 2005, Goetz et al., 2010, Goetz et al., 2012b) and mouse βKlotho (Goetz et al., 2012a) proteins from insect or mammalian cells, while others have resorted to protein fusions to potentially solubilising modules such as human IgG1 Fc, expressed and purified from the conditioned media of CHO cells (Tohyama et al., 2004). Extracellular and full length FGFRs have also been previously expressed and purified from insect and mammalian cells by some commercial labs.

In order to produce the large amounts of each recombinant protein required for this work, a panel of constructs was initially designed for expression testing in *E. coli* (Table 2.3). The basic design of these constructs is shown in Figure 4.9. Each endocrine FGF and βKlotho construct was synthetically codon-optimised for expression in *E. coli* and

comprised a PreScission™ protease-cleavable C-terminal 2StrepII-10His tag. Construct boundaries for FGFs were defined as the full length mature protein sequences without the N-terminal secretion signal peptide, whereas construct boundaries for βKlotho domains were determined based on the analyses in section 4.3.1. All constructs were also generated with and without a non-native N-terminal secretion signal peptide from *E. caratovora* pectate lyase B (pelB) or human serum albumin (HSA) proteins to test for periplasmic expression. The two extracellular FGFR1c constructs were generated using standard molecular biology subcloning techniques from a native human full length FGFR1c mammalian cDNA construct which was a gift (AstraZeneca). This construct comprised a C-terminal FLAG tag, but both subcloned extracellular FGFR1c constructs were engineered with the PreScission™ protease-cleavable C-terminal 2StrepII-10His tag. Construct boundaries were defined as either the full length extracellular region D1-D3 without the N-terminal signal peptide, or the previously characterised D2-D3 domain construct (Plotnikov et al., 1999). Attempts were also made to subclone some of these FGFR1c and βKlotho constructs into the pTriEx™4 vector (Novagen) using a variety of different molecular biology techniques but this proved to be challenging and only a few constructs could be successfully generated (Table 2.3). For full methods see sections 2.1 and 2.2.2.

These constructs were then subjected to a series of small scale (2-5 mL) expression tests in *E. coli* C41 (DE3) cells. Additionally, *E. coli* XL10 Gold cells were employed to test for periplasmic expression based on a previously published protocol (Hage et al., 2015). A variety of expression conditions were tested including different temperatures and durations of expression. Harvested cell samples were lysed and separated into soluble and insoluble fractions and analysed by SDS-PAGE and immunoblotting. Only a few constructs expressed reproducibly and these successful protein expressions are shown in Figure 4.10. The human βKlotho KL1 domain expressed well, both in the soluble and insoluble fractions, with increasing soluble expression upon lowering of the expression temperature down to 15°C (Figure 4.10a, left panels). The KL2 domain expressed best at 37°C and in inclusion bodies, with only weak soluble expression at lower temperatures (Figure 4.10a, right panels). The longer KL1-KL2 protein only expressed at 37°C and in inclusion bodies (Figure 4.10c). FGF19 and FGF21 and both lengths of extracellular FGFR1c expressed in inclusion bodies (Figure 4.10b and 4.10d), as expected from the literature (mentioned above in section 4.3.2). Expression of FGF19 was reproducibly lower than FGF21 expression, similar to the expression pattern seen for the longer FGFR1c D1-D3 (22-374) compared to the shorter D2-D3 (142-365) construct. No periplasmic expression was observed for any of the FGF, FGFR or βKlotho constructs in either of the two *E. coli* strains tested (data not shown). Additionally, some

other βKlotho constructs expressed in one expression trial but could not be reproduced and so the SDS-PAGE gels and immunoblots are not shown.

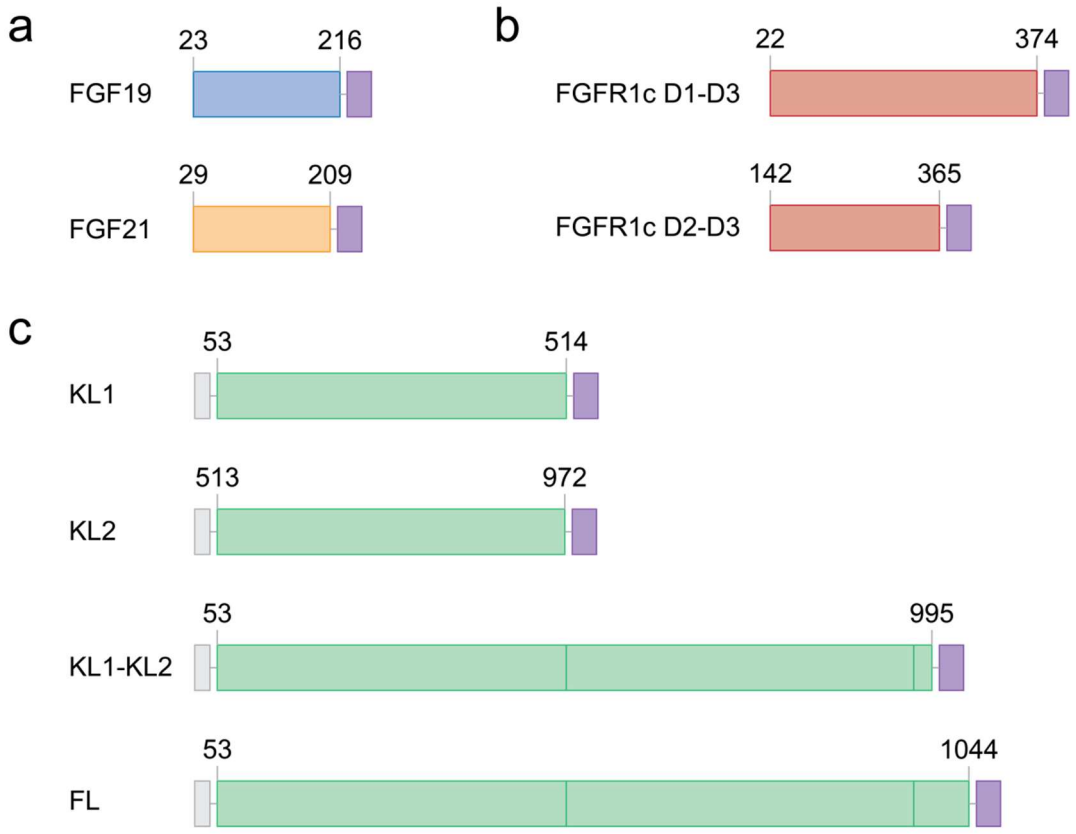


Figure 4.9 Bacterial expression constructs: design and nomenclature. Fundamental design of human endocrine FGF, FGFR1c and βKlotho constructs designed for *E. coli* expression. All constructs were codon optimised for expression in *E. coli* and comprised a PreScission™ protease-cleavable C-terminal 2StrepII-10His tag (light purple). All βKlotho constructs were also generated with and without a non-native N-terminal secretion signal peptide (light grey) from *E. caratovora* pectate lyase B (pelB) or human serum albumin (HSA) to test for periplasmic expression. **(a)** FGF19 (23-216) and FGF21 (29-209) construct boundaries were defined as the full length mature protein sequences without the N-terminal secretion signal peptide. **(b)** FGFR1c extracellular domain constructs were prepared as either the full length extracellular region 22-374 (D1-D3) without the N-terminal signal peptide or the shorter length 142-365 (D2-D3) construct which has been used in the literature by several groups since its first use (Plotnikov et al., 1999) and comprises the minimal paracrine FGF ligand binding region of FGFRs. **(c)** βKlotho construct boundaries were defined based on the bioinformatics and manual literature mining analyses performed in this work (Figures 4.7, 4.8 and Table 4.4) in section 4.3.1.

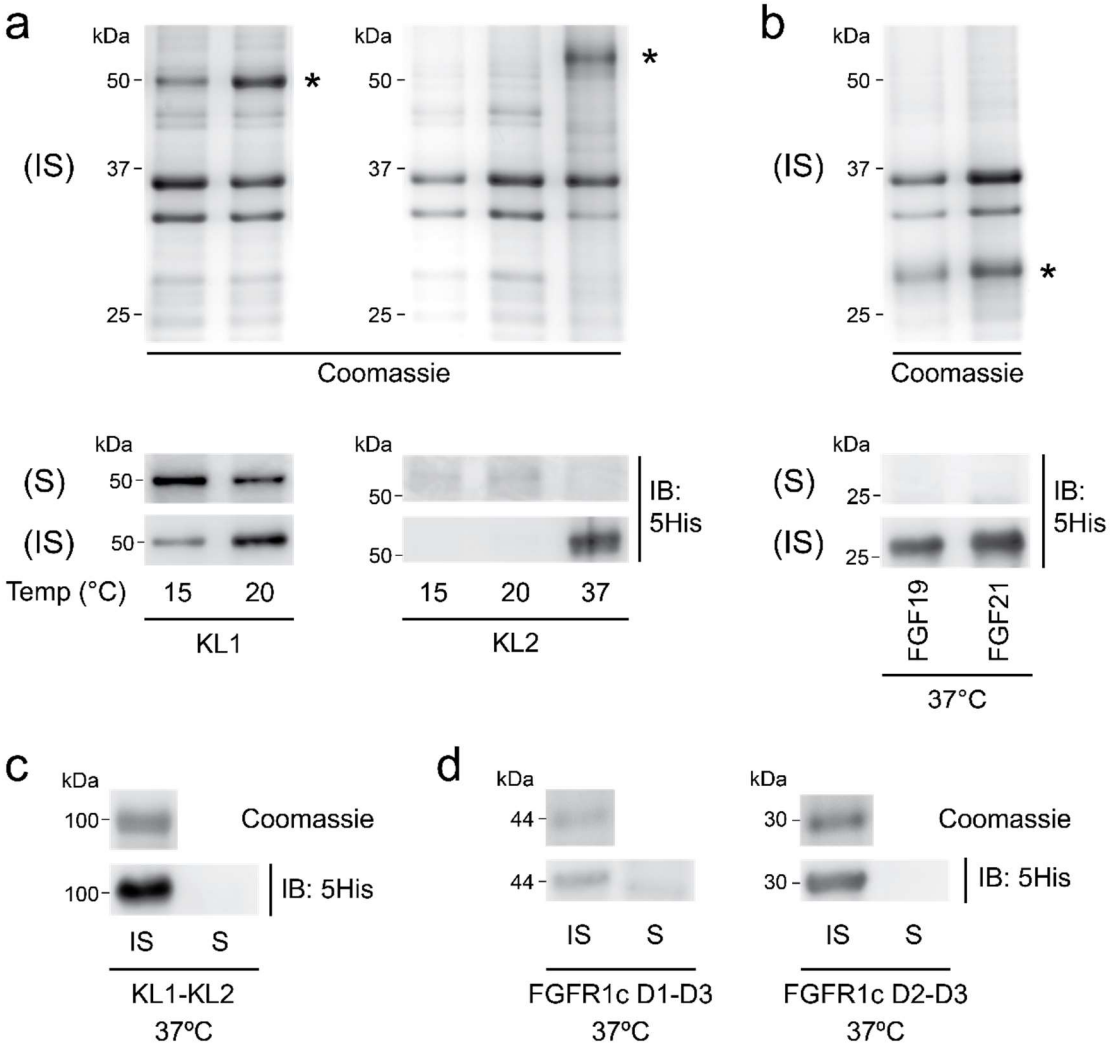


Figure 4.10 Successful recombinant protein expression from bacterial expression tests in *E. coli*. From the panel of bacterial expression constructs generated (Table 2.3 and Figure 4.9) those that expressed in *E. coli* reproducibly in either the soluble (S) or insoluble (IS) fractions from overnight expressions at different temperatures are shown here as identified by immunoblotting with the indicated anti-5His antibody. The corresponding Coomassie stained SDS-PAGE gels of IS fractions for which the recombinant proteins could be easily identified on the gels (proteins were normally indistinguishable on gels of S fractions due to the presence of too many proteins at the same molecular weights in crude lysates) are also shown for all constructs (upper panels). The correct bands on the SDS-PAGE gels corresponding to bands of recombinant proteins identified by immunoblotting are marked by asterisks in **a** and **b**. Gels and immunoblots shown are representative of at least three expression trials for each construct. For full experimental details see section 2.3.3. **(a)** Soluble expression of the βKlotho KL1 domain decreased when the expression temperature was increased. Very low levels of soluble KL2 domain could be detected at lower expression temperatures, but expression increased at higher temperatures with protein instead expressing in inclusion bodies in the IS fraction. **(b)** Both endocrine FGFs expressed only in the IS fraction, as expected from the literature (see section 4.3.2). **(c)** Expression of the entire extracellular region of βKlotho (comprising KL1-KL2 domains, also known as soluble Klotho) could only be detected in the IS fraction. **(d)** The longer length FGFR1c protein (D1-D3) expressed less well than the shorter length (D2-D3) construct, and both expressed only in inclusion bodies in the respective IS fractions.

The KL1 and KL2 domain protein expressions were scaled up to test purification from the soluble fractions (using optimal expression temperatures as determined from the expression tests, Figure 4.10a) and in parallel to attempt refolding of the inclusion bodies. Soluble lysate was purified by affinity and ion exchange chromatography followed by analytical SEC. Representative chromatograms and gels are shown in Figure 4.11. Both KL1 and KL2 eluted in the void volume as soluble aggregates as a minor species alongside various contaminating proteins (Figure 4.11a, upper panel, and 4.11b, two left panels). It was not possible to dissociate these aggregates in the presence of high salt or reducing agent (Figure 4.11a, lower panel and 4.11b, two right panels). It was decided that this was therefore not a practically feasible method to produce recombinant KL domain proteins.

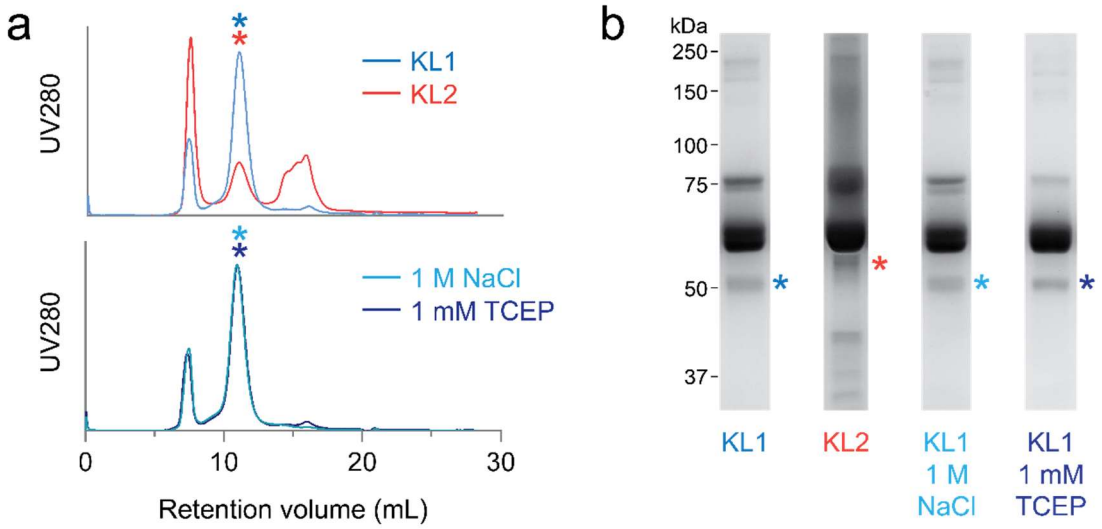


Figure 4.11 Purification of individual βKlotho KL1 and KL2 domain proteins from soluble expressions in *E. coli*. (a) Size exclusion chromatography (SEC) chromatograms for the final stage chromatographic purification of recombinant KL1 and KL2 domain proteins under normal salt (137 mM) and non-reducing buffer conditions (upper panel), and in the presence of high salt (1 M NaCl) or reducing agent (1 mM TCEP) for KL1 from a single experiment (lower panel). The asterisks label specific peaks and highlight KL protein eluting in the void volume under all buffer conditions tested. **(b)** SDS-PAGE analysis of proteins eluting in the peaks marked by asterisks in **a**, demonstrating the (lack of) purity of KL proteins. On the Coomassie stained gels, bands corresponding to KL proteins (based on similar results from pilot experiments, see Figure 4.10a) are indicated by matching asterisks to those in **a**. Chromatograms and gels shown are representative of five and four independent large scale purifications performed with soluble *E. coli* extracts of KL1 and KL2 respectively.

All inclusion body samples were washed sequentially with three different wash buffers prior to denaturation and refolding via several different workflows (Figure 4.12a-b). Inclusion body denaturation using 8 M urea was initially attempted but subsequently 6 M guanidine hydrochloride was used as the denaturant for all experiments due to its ease of use compared to urea, the latter of which is prone to crystallisation, particularly at high concentrations. Attempts to refold inclusion bodies for each KL domain protein led to precipitation of all protein during various stages of the refolding procedures (data not shown). A similar outcome emerged from scaling up and attempting to refold the KL1-KL2 protein from inclusion bodes (Figure 4.12c). In the case of the FGFR1c proteins, refolding of the inclusion bodies appeared to be initially successful up to the stage of the Ni^{2+}-NTA affinity chromatography step but all protein degraded during dialysis prior to the final SEC step (Figure 4.12d). The refolding conditions tested for all FGFR1c and βKlotho constructs in each case are detailed in Table 4.5. Due to the challenging and empirical nature of purifying proteins via inclusion body refolding, further experiments were not continued via such methods for these relatively large (for refolding) extracellular domain proteins following these unsuccessful refolding trials.

Table 4.5 Specific refolding conditions tested for βKlotho and FGFR1c inclusion body (IB) preparations in this work. All refolding solutions also contained 2 mM L-cysteine and 0.2 mM L-cystine (disulphide shuffling agents). It should be noted that dilute, low denaturing concentrations of guanidine hydrochloride and/or urea (from previous steps of IB denaturation and solubilisation) were also present in the final refolding solutions. CHAPS: 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulphonate, PPS: 3-(1-pyridinio)-1-propanesulphonate.

Protein	Buffer [mM]	pH	NaCl [mM]	Refolding additive	EDTA [mM]	DTT [mM]
KL1	Tris/ethanolamine [50]	8.0	0	PPS [1 M]	0.1	0
KL2	Tris/ethanolamine [50]	8.0	0	PPS [1 M]	0.1	0
	Tris [50]	8.5	100	L-arginine [0.5 M]	0	0
KL1-KL2	Tris [100]	8.0	100	PPS [1 M]	0.1	0
	Tris [100]	8.0	100	L-arginine [0.5 M]	0.1	0
	Tris [100]	8.0	100	CHAPS [80 mM]	0.1	0
FGFR1c D1-D3	HEPES [25]	7.5	150	Glycerol [10%]	0.25	1
	Tris [100]	8.5	100	L-arginine [0.5 M]	0.25	1
	Tris/ethanolamine [50]	8.0	0	PPS [1 M]	0.25	1
	Tris/ethanolamine [50]	7.0	0	PPS [1 M]	0.25	1
FGFR1c D2-D3	HEPES [25]	7.5	150	Glycerol [10%]	0.25	1
	Tris [100]	8.5	100	L-arginine [0.5 M]	0.25	1
	Tris/ethanolamine [50]	8.0	0	PPS [1 M]	0.25	1
	Tris/ethanolamine [50]	7.0	0	PPS [1 M]	0.25	1

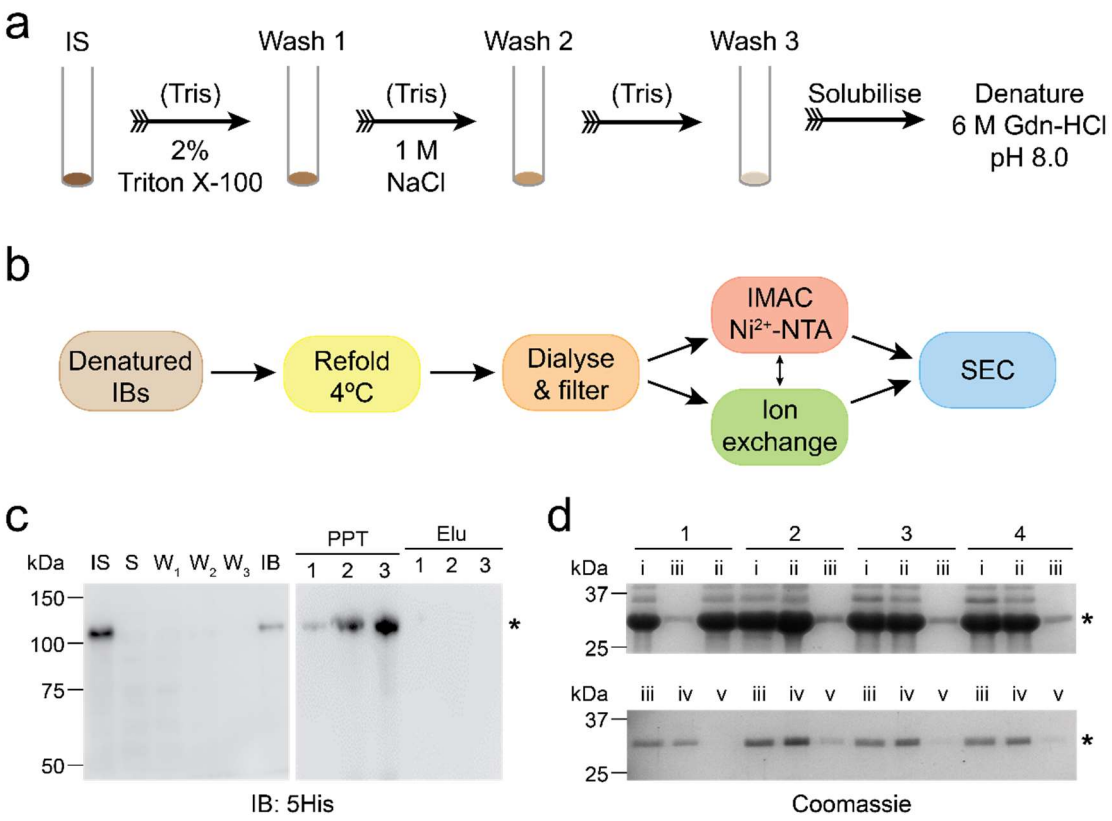


Figure 4.12 Inclusion body (IB) preparation, general refolding workflow and preliminary refolding trials for βKlotho and FGFR1c. All IBs expressed in *E. coli* were treated in the same way prior to denaturation and refolding. **(a)** The insoluble (IS) fraction (dark brown) containing IBs was sequentially washed with three different buffers to remove different protein and nucleic acid contaminants. Washed IBs (light brown) were then solubilised and denatured using guanidine hydrochloride (Gdn-HCl) prior to refolding. **(b)** Schematic representation of the general refolding workflow used for all protein IBs. All parameters were empirically determined based on observations made during refolding trials. Flexibility in the workflow was crucial for customising the method for individual proteins based on variables such as temperature, pH, ionicity and viscosity of the refolding solution while trying to minimise protein aggregation and precipitation. For full details of refolding conditions tested for βKlotho and FGFR1c, see Table 4.4, and for refolding and purification of FGF19 and FGF21 see Figures 4.13 and 4.14. **(c)** Immunoblots with the indicated anti-5His antibody showing the progression of the KL1-KL2 protein through the refolding workflow in **a** (left panel) and following refolding, dialysis, filtration and Ni^{2+}-NTA affinity chromatography (right panel) as marked by the single asterisk. All KL1-KL2 protein precipitated during the refolding step prior to Ni^{2+}-NTA affinity chromatography, under all three (1-3) refolding conditions tested (refolding condition details in Table 4.5). IS – insoluble fraction following *E. coli* lysate clarification, S – soluble lysate following *E. coli* lysate clarification, W_1, W_2, W_3 – washes 1-3 as in **a**, IB – washed IBs prior to solubilisation/denaturation with Gdn-HCl, PPT – refolding precipitate removed by filtration following dialysis and prior to Ni^{2+}-NTA affinity chromatography, Elu – elution fractions following Ni^{2+}-NTA affinity chromatography. **(d)** Coomassie stained gels representative of at least two independent refolding trials performed for FGFR1c D2-D3 (142-365) under four (1-4) refolding conditions (details in Table 4.5) demonstrating partially successful purification of these proteins (marked by a single asterisk) following the Ni^{2+}-NTA affinity chromatography step. All protein then degraded during dialysis prior to the size exclusion chromatography (SEC) step. A similar result was obtained with FGFR1c D1-D3 (22-374) which expressed more poorly than the D2-D3 construct, so this data is not shown. i – refolded protein prior to filtration, ii – refolded and filtered protein prior to Ni^{2+}-NTA affinity chromatography, iii – elution fractions following Ni^{2+}-NTA affinity chromatography, iv – dialysed Ni^{2+}-NTA elution fractions prior to SEC (not filtered to remove precipitate), v – dialysed and filtered Ni^{2+}-NTA elution fractions prior to SEC (precipitate removed).

4.3.3 Refolding of endocrine FGFs from inclusion bodies and quality control

Of all the proteins expressed in inclusion bodies, two which were successfully refolded to produce soluble protein were FGF19 and FGF21. A refolding method adapted from one used previously to refold FGF19 (Harmer et al., 2004) was employed for both FGF19 and FGF21. Inclusion bodies were first washed as described above (Figure 4.12a) and then subjected to several different refolding trials in parallel (Figure 4.13a). Optimal yields of the FGFs required one day under refolding conditions with L-cysteine/L-cystine as the disulphide shuffling agents, and overall FGF21 yielded larger amounts of protein than FGF19 (Figure 4.13b). The immunoblot in Figure 4.13b also suggested that extra days under refolding conditions did not appear to provide any advantage or disadvantage on final yields of FGF21, however FGF19 appeared to be less stable and more was lost through degradation when longer refolding durations were used. The affinity purification step also clearly provided the most efficient separation of FGFs from all other proteins in the refolding solution (Figure 4.13c). In contrast, ion exchange chromatography did not add any chromatographic benefit to the purification and was therefore eliminated from the optimised protocol (data not shown).

From all the various purification 'routes' tested in the refolding workflow, the best method for both FGFs (that which ultimately produced the largest quantities of refolded, soluble protein) is highlighted in Figure 4.13a and detailed in Figure 4.14b. Expression of each FGF was then scaled up (Figure 4.14a) and the protein refolded and purified using this method. Representative SEC chromatograms and Coomassie stained SDS-PAGE gels demonstrating protein purity are shown in Figure 4.14c. Peaks on the chromatograms marked with a single asterisk comprised monomeric, soluble FGFs. Peaks marked with double asterisks comprised FGFs in soluble dimers or aggregates as well as various other protein contaminants. Protein in fractions from both monomer and aggregate peaks was separately retained, concentrated and stored until it was required for experiments. Figure 4.14c (upper and lower right panels) and 4.14f show the resulting purity of the concentrated proteins.

In addition to the two endocrine FGFs, a full length human FGF1 wildtype protein was also produced as a control paracrine FGF for further experiments in which FGF19 and FGF21 were used. This protein was expressed in *E. coli* where it expressed abundantly in the soluble fraction (unlike FGF19 and FGF21 which only expressed in inclusion bodies in the insoluble fraction) and so it was purified from the soluble lysate. The FGF1 construct design is shown in Figure 4.14d. A single peak was observed on the SEC chromatogram and this contained pure, monomeric FGF1 protein, demonstrated using SDS-PAGE analysis and immunoblotting (Figure 4.14e-f).

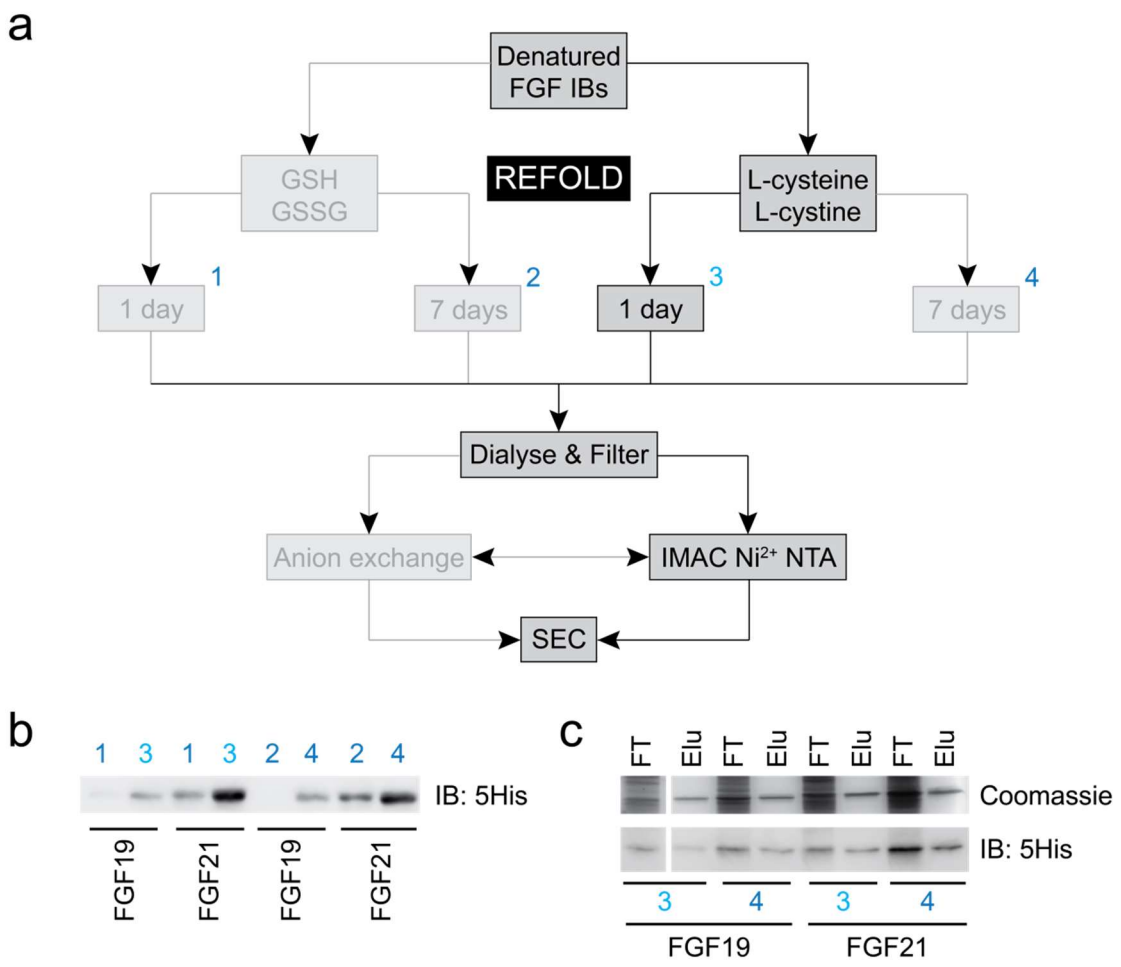


Figure 4.13 Refolding trials for FGF19 and FGF21 expressed in inclusion bodies (IBs) in *E. coli*. (a) Schematic representation of the workflow used to test refolding conditions for recombinant FGF19 and FGF21 IBs based on a published protocol (Harmer et al., 2004). GSH and GSSG (reduced and oxidised glutathione) or L-cysteine and L-cystine (reduced and oxidised L-cysteine) were tested as disulphide shuffling agents. Refolding samples were taken after 1 day or 7 days under the different refolding conditions (condition 'routes' numbered 1-4). Various chromatographic steps were tested for the purification, in different orders post-refolding. The method found to be optimal for both FGFs is highlighted in the flow diagram (condition 3). **(b)** Immunoblot with the indicated anti-5His antibody showing relative soluble yields of FGF19 and FGF21 under various refolding conditions tested (numbered as in **a**), representative for two independent refolding trials performed for all 'routes' in the workflow shown. **(c)** Coomassie stained SDS-PAGE gel (upper panel) and corresponding immunoblot with the indicated anti-5His antibody (lower panel) showing the efficient separation of His-tagged recombinant FGF19 and FGF21 from the refolding solution using affinity (Ni^{2+}-NTA) chromatography. FT and Elu refer to the flowthrough and eluate fractions after proteins from the indicated refolding 'routes' (numbered as in **a**) were subjected to this chromatographic step. For full experimental details see section 2.6.4.

In order to assess the bioactivity of the refolded recombinant FGFs, a series of cell-based activity assays were set up. First, assays were set up to observe the effects of FGFs on downstream signalling in 3T3-L1 mouse adipocytes. The adipocytes can be generated *in vitro* from 3T3-L1 pre-adipocyte precursor cells using an established differentiation protocol (Zebisch et al., 2012). These adipocytes are known to express both FGFR1 and βKlotho upon differentiation, and have been shown to respond to both FGF19 and FGF21 in a dose dependent manner (Kurosu et al., 2007, Ogawa et al., 2007, Yang et al., 2012). For full experimental details see section 2.12. In a first experiment, cells were treated with a single low concentration of each recombinant endocrine FGF and FGF1 (control). In a second experiment, mini dose response assays were set up whereby cells were treated with higher concentrations of each recombinant FGF or additionally with commercially available FGF19 and FGF21 (R&D Systems) as further controls. Cellular responses to FGFs were recorded in each case by monitoring the level of ERK phosphorylation (Figure 4.15a and 4.15c). Concurrently, the differentiation process of 3T3-L1 pre-adipocytes to mature adipocytes was also followed by observing increased ERK phosphorylation over the course of the differentiation process (Figure 4.15b). In both the first and second assays the cells responded to FGF1 as depicted by elevated ERK phosphorylation, although the dose response was not very apparent at the highest concentration of FGF1 tested. In both experiments, there was a small increase in ERK phosphorylation for both FGF19 and FGF21 but the responses were not very clear and there was no obvious dose response for both the refolded FGFs and surprisingly also for the control commercially acquired endocrine FGFs. This assay was also repeated with lower concentrations of FGFs but produced similarly disappointing results (data not shown).

It was therefore decided to apply circular dichroism (CD) spectroscopy to determine whether the recombinant FGFs were correctly folded. This useful biophysical technique can provide information about the secondary structural content (and therefore the 'foldedness') of proteins, but cannot report on oligomerisation status or bioactivity. Despite this, validation of the 'foldedness' of the FGFs following *in vitro* refolding was considered the minimal requirement to proceed with the FGFs in further experiments. Based on the analysis of existing atomic structures of human FGF1 (more than 70 structures in the RCSB PDB, e.g. PDB: 1RG8 and 2ERM) and FGF19 (PDB: 1PWA and 2P23), these FGFs are expected to be composed of predominantly β-strands and unstructured loop regions, with little 3_{10} and α-helical content (Figure 4.16). Therefore, if the recombinant FGFs were correctly folded, the resulting CD spectra would be expected to display a strong negative peak *between* the wavelengths at which the primary negative peaks for β-strands (218 nm) and unstructured regions (195 nm) would be expected.

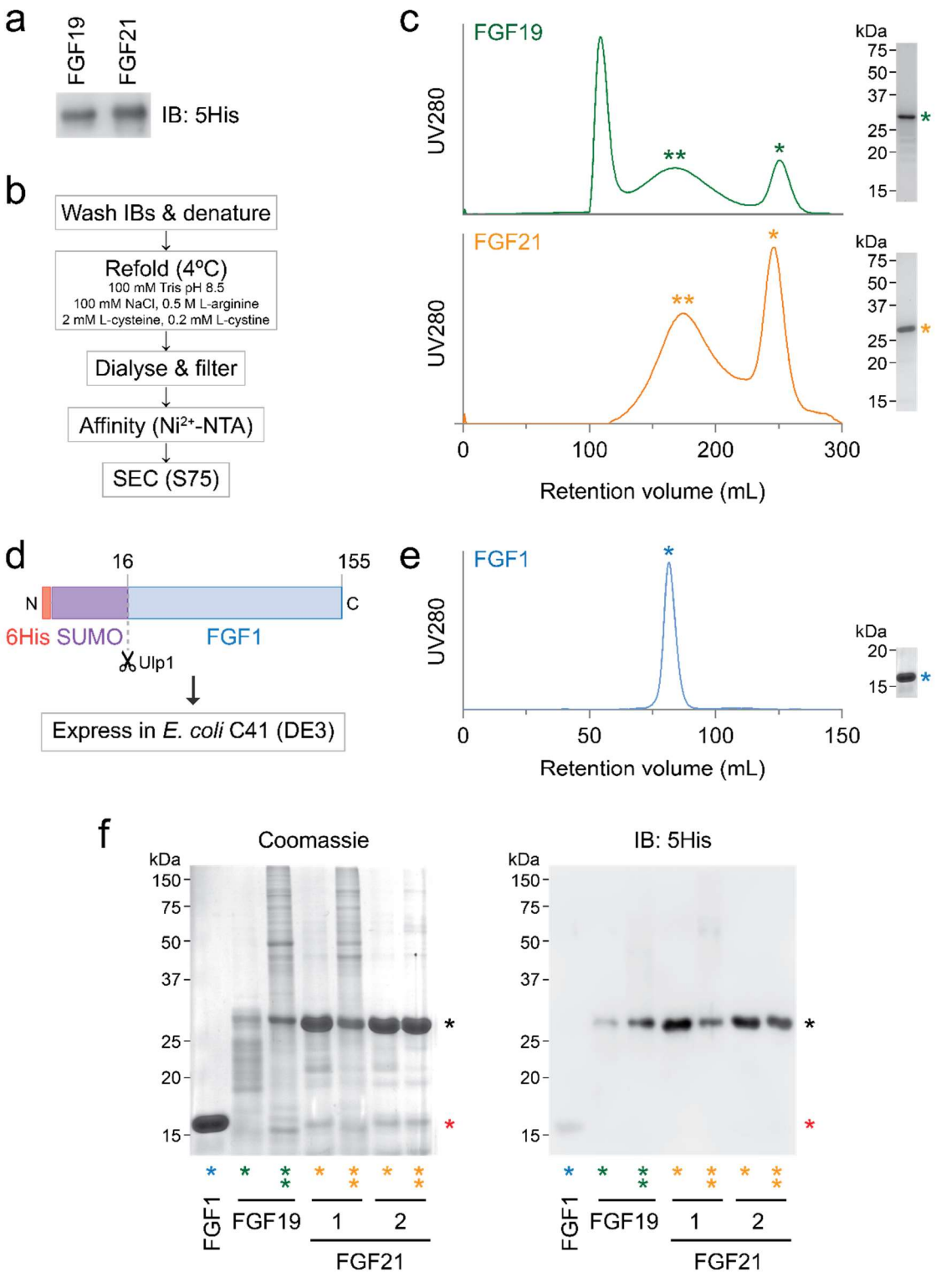

a
FGF19
FGF21
IB: 5His
b
Wash IBs & denature
Refold (4°C)
100 mM Tris pH 8.5
100 mM NaCl, 0.5 M L-arginine
2 mM L-cysteine, 0.2 mM L-cystine
Dialyse & filter
Affinity (Ni²⁺-NTA)
SEC (S75)
c
FGF19
FGF21
UV280
kDa
75
50
37
25
20
15
0
100
200
300
Retention volume (mL)
d
16
155
N
C
6His SUMO
FGF1
Ulp1
Express in E. coli C41 (DE3)
e
FGF1
UV280
kDa
20
15
0
50
100
150
Retention volume (mL)
f
Coomassie
IB: 5His
kDa
150
75
50
37
25
20
15
FGF1
FGF19
1
2
FGF21

Figure 4.14 Purification of refolded FGF19, FGF21 and FGF1 proteins. (a) Immunoblot using the indicated anti-5His antibody showing the relative expression of FGF19 and FGF21 in bacterial inclusion bodies. **(b)** Details of the final conditions and method developed for refolding of recombinant FGF19 and FGF21 based on refolding trials in Figure 4.13. **(c)** SEC chromatograms from the final stage of purification of FGF19 (upper left panel) and FGF21 (lower left panel) representative for one and three independent completed protein preparations of FGF19 and FGF21 respectively. The single asterisk highlights peaks containing monomeric, soluble FGFs, and adjacent Coomassie stained SDS-PAGE gels (upper and lower right panels) show the purity of the monomeric proteins. Double asterisks highlight peaks containing FGF protein aggregates as well as various other contaminants. Only monomeric protein preparations were used in further experiments in this work. **(d)** Schematic diagram showing construct design for the full length recombinant human FGF1 (16-155) protein expressed in this work. The plasmid construct was provided by Tom Bunney. Recombinant FGF1 protein expressed in *E. coli* using this construct was found in the soluble fraction (data not shown). **(e)** SEC chromatogram for the final stage of purification of FGF1 (left panel). The protein eluted in a single peak (marked by a single asterisk) which contained monomeric protein. SDS-PAGE analysis (right panel) shows the purity of the monomeric FGF1 used in further experiments in this work. **(f)** Coomassie stained SDS-PAGE gel (left panel) and corresponding immunoblot with the indicated anti-5His antibody (right panel) demonstrating the level of protein purity and degradation observed in purified preparations of recombinant FGF1, refolded FGF19 and two independently refolded FGF21 preparations (1 and 2). Monomer (single asterisk) and aggregate (double asterisks) fractions are shown following variable durations of storage of concentrated purified proteins at -80°C. Bands corresponding to FGF19 and FGF21 (black asterisk) and FGF1 (red asterisk) are indicated adjacent to the SDS-PAGE gel and immunoblot images. For full experimental details, see sections 2.6.2 and 2.6.4.

CD scans were measured for each recombinant FGF monomeric protein (protein from gels and SEC chromatogram peaks marked with a single asterisk in Figure 4.14c and 4.14e-f) between 180-300 nm wavelengths. Two refolded and purified samples of independently prepared FGF21 proteins were included in the experiment, where they are referred to as 'FGF21' and 'FGF21 (2)'. Data were processed and analysed using standard CD spectral analysis software and two independent online analysis algorithms (for full methods see section 2.20.2). Both analysis methods produced very similar final results, so the best quality data between 190-240 nm wavelengths from the better of the two analyses are displayed in Figure 4.17a. All four samples of FGFs displayed a strong negative peak between 203-206 nm, characteristic of proteins comprising largely a mixture of parallel and anti-parallel β-sheet and loop, turn or unstructured regions. The small differences in the nature of the negative peak for the two independently prepared FGF21 samples suggest an inherent susceptibility to conformational instability over different durations of long-term storage of the two protein preparations at -80°C. Overall, the CD traces compared well with secondary structural predictions from the published FGF atomic structures (Figure 4.16) and with previously published CD traces measured for FGF1 and FGF21 (Xu et al., 2012). Based on these results, the recombinant FGFs prepared in this work were deemed folded and of sufficient quality for use in further experiments.

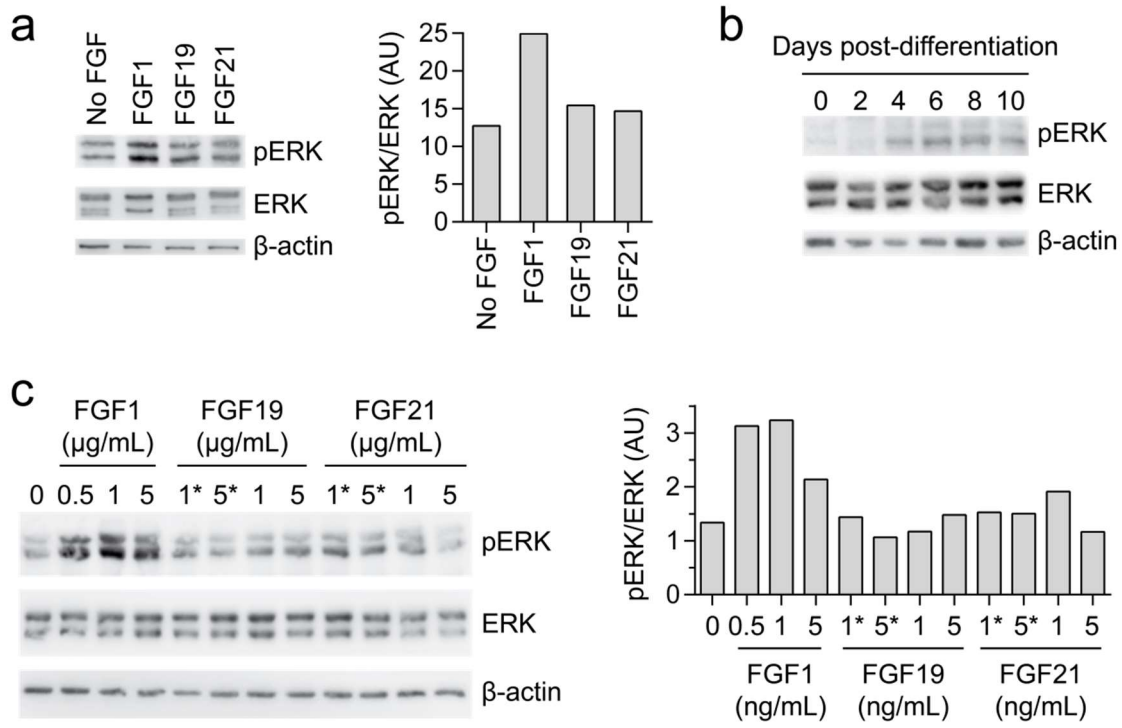


Figure 4.15 Bioactivity assays for refolded FGF proteins. (a) All FGFs activate ERK1/2 signalling in 3T3-L1 adipocytes, to differing degrees. Differentiated 3T3-L1 adipocytes were stimulated with recombinant FGF1, FGF19 and FGF21 (80 ng/mL) for 10 minutes before the cells were harvested, lysed and analysed for activation of downstream signalling pathways. Immunoblots with the indicated antibodies representative of two repeats from a single experiment are shown in the left panel. The results from densitometric analyses performed on these data are displayed in the right panel. **(b)** Increases in ERK1/2 signalling upon differentiation of 3T3-L1 pre-adipocytes to adipocytes. Differentiation was induced on Day 0, and cells were harvested and lysed on the indicated day post-induction of differentiation. Lysates were then analysed for activation of downstream signalling pathways using the indicated antibodies and the resulting immunoblots are shown. **(c)** An experiment similar to that in **a** was performed using higher concentrations of all recombinant FGFs as well as commercially obtained FGF19 and FGF21 (R&D Systems, indicated by single asterisks). Resulting immunoblots are shown (left panel) and results from the densitometric analyses performed on these data are presented (right panel). For full experimental details, see sections 2.12 and 2.14.

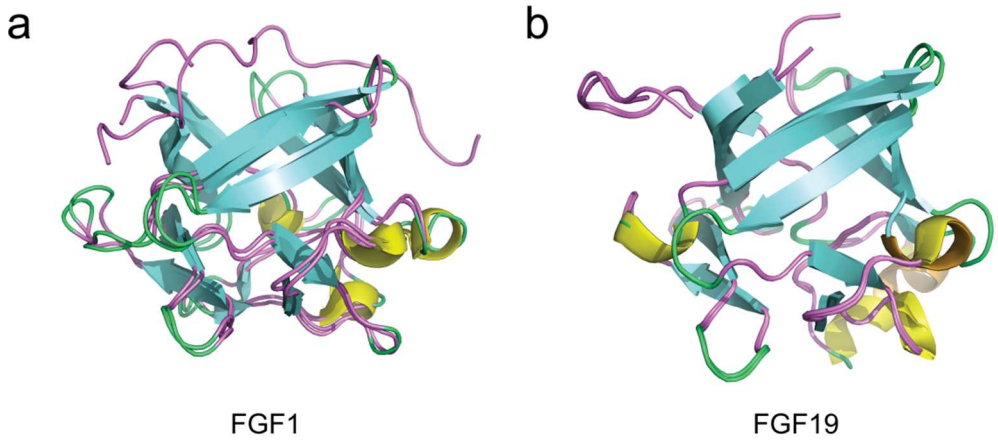
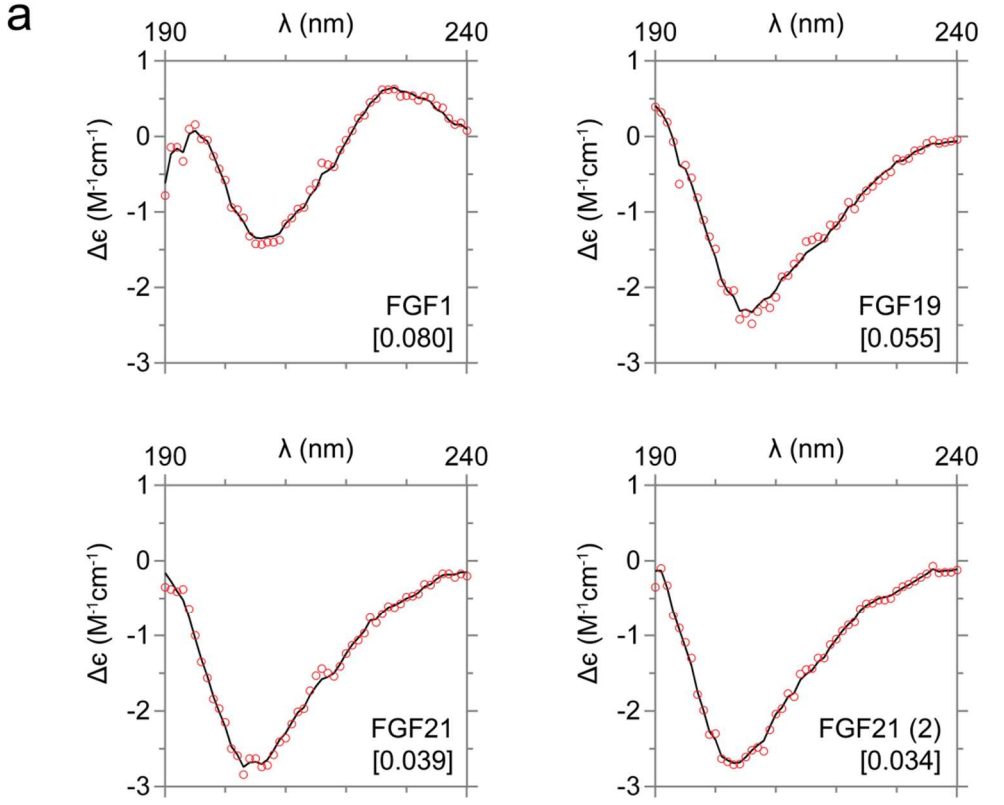


Figure 4.16 Comparison of secondary structural composition of FGF1 and FGF19. Overlays of two crystal structures from the RCSB PDB for each of **(a)** human FGF1 [PDB: 1RG8 and 2ERM] (Bernett et al., 2004, Canales et al., 2006) and **(b)** human FGF19 [PDB: 1PWA and 2P23] (Harmer et al., 2004, Goetz et al., 2007) shown in cartoon representation and in the same orientation. The atomic models are coloured by secondary structure with α-helices (orange), 3_{10} helices (yellow), β-strands (blue), turns (green) and loops (purple). The FGFs are composed of predominantly β-strands and unstructured loop regions, with little 3_{10} and α-helical content. The structures of endocrine FGF19 retain the conserved core β-trefoil fold observed in all paracrine FGF structures including FGF1.


a
190
λ (nm)
240
Δε (M-1cm-1)
1
0
-1
-2
-3
FGF1
[0.080]
FGF19
[0.055]
FGF21
[0.039]
FGF21 (2)
[0.034]

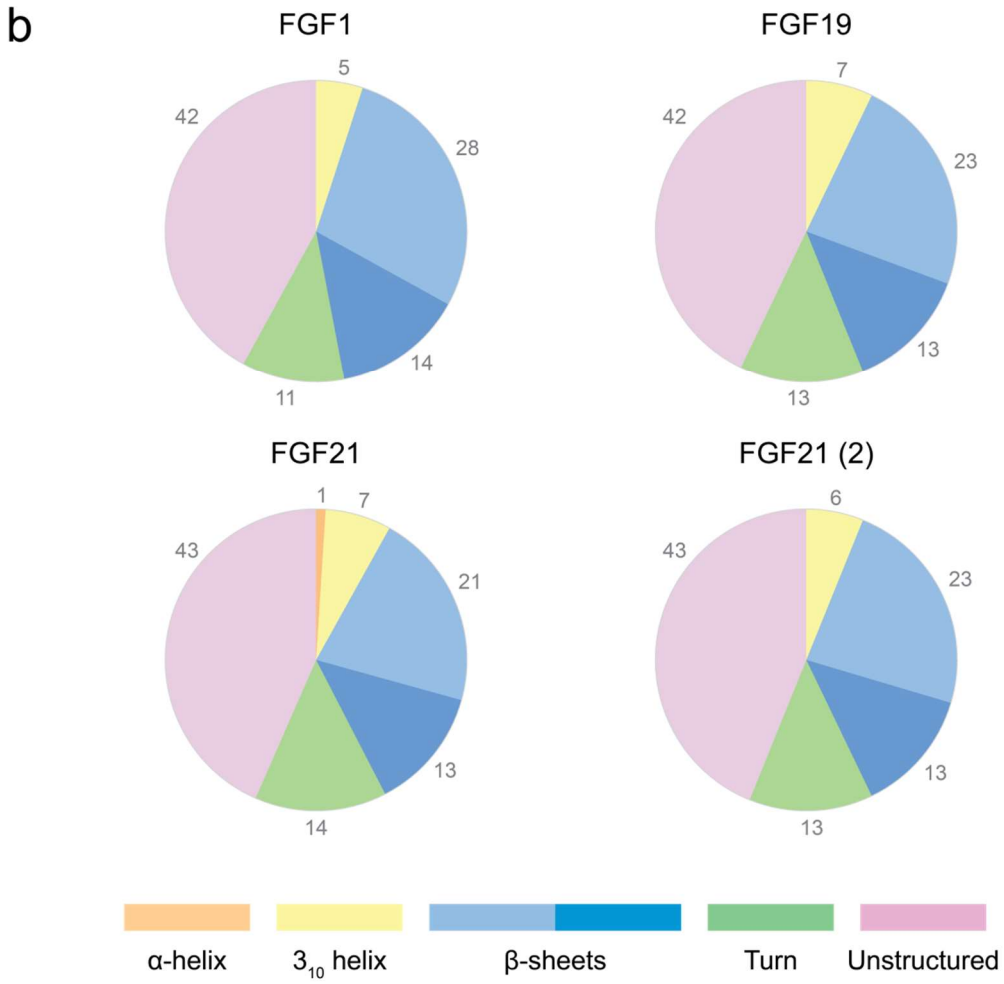

b
FGF1
5
28
14
11
42
FGF19
7
23
13
13
42
FGF21
1
7
21
13
14
43
FGF21 (2)
6
23
13
13
43
α-helix
3₁₀ helix
β-sheets
Turn
Unstructured

Figure 4.17 Recombinant FGFs are folded and comprise primarily β-strands and unstructured / loop regions. Circular dichroism (CD) spectroscopy was employed to assess the foldedness of the recombinant FGFs. In each experiment, ten scans for each protein between 180-300 nm wavelengths, were measured, averaged, reference subtracted and normalised using CDtool spectral analysis software (Lees et al., 2004). Some proteins were also scanned in two duplicate experiments. Processed data were analysed by multiple algorithms and the highest quality results from CDSSTR (Manavalan and Johnson, 1987) with reference dataset SP175 (Lees et al., 2006) between 190-240 nm wavelengths are displayed. **(a)** Processed CD scans for each recombinant FGF showing experimentally measured values (open red circles) and non-linear fits (continuous black lines) between 190-240 nm wavelengths. NRMSD values (Mao et al., 1982) for the fits are displayed in square brackets, showing that data fits for all CD scans are robust. **(b)** Proportional secondary structure assignment for each FGF performed using the same algorithm as in **a**. Specific secondary structure content percentages are annotated around each pie chart. For full experimental details, see section 2.20.2.

4.3.4 Redesigning FGFR1c and βKlotho constructs for mammalian expression

As described in section 4.3.2, it was not possible to successfully purify extracellular domain constructs of any length tested of FGFR1c or βKlotho from expression in *E. coli*. Extracellular FGFR1c and βKlotho proteins have, however, been overexpressed and purified previously from mammalian cells by some commercial labs and academic groups, such as FGFR1 from Sf9 insect and HEK 293 cells (OriGene), human βKlotho from CHO cells (Yie et al., 2009) and mouse βKlotho from HEK 293 cells (Goetz et al., 2012a). Therefore the FGFR1 and βKlotho expression constructs were next redesigned to test for overexpression in mammalian cell lines. All existing constructs were first codon-optimised for expression in mammalian cell systems. Non-membrane bound extracellular constructs were designed as fused to either the native (βKlotho) or a non-native (human CD33) N-terminal secretion signal peptide. The construct boundaries were maintained for all FGFR1c and βKlotho constructs generated but this time the focus was on the longer βKlotho constructs (KL1-KL2 and full length) rather than the shorter length individual KL domain constructs. In addition, a full length native mouse βKlotho cDNA construct with a C-terminal fusion to eGFP (mβe) was obtained (Addgene) to add to the expression panel. This mouse construct was modified by the addition of an engineered protease cleavable C-terminal 6His tag to aid purification (mβTe6H). The designs of all mammalian expression constructs generated for this work are listed in Table 2.4 and displayed in Figure 4.18.

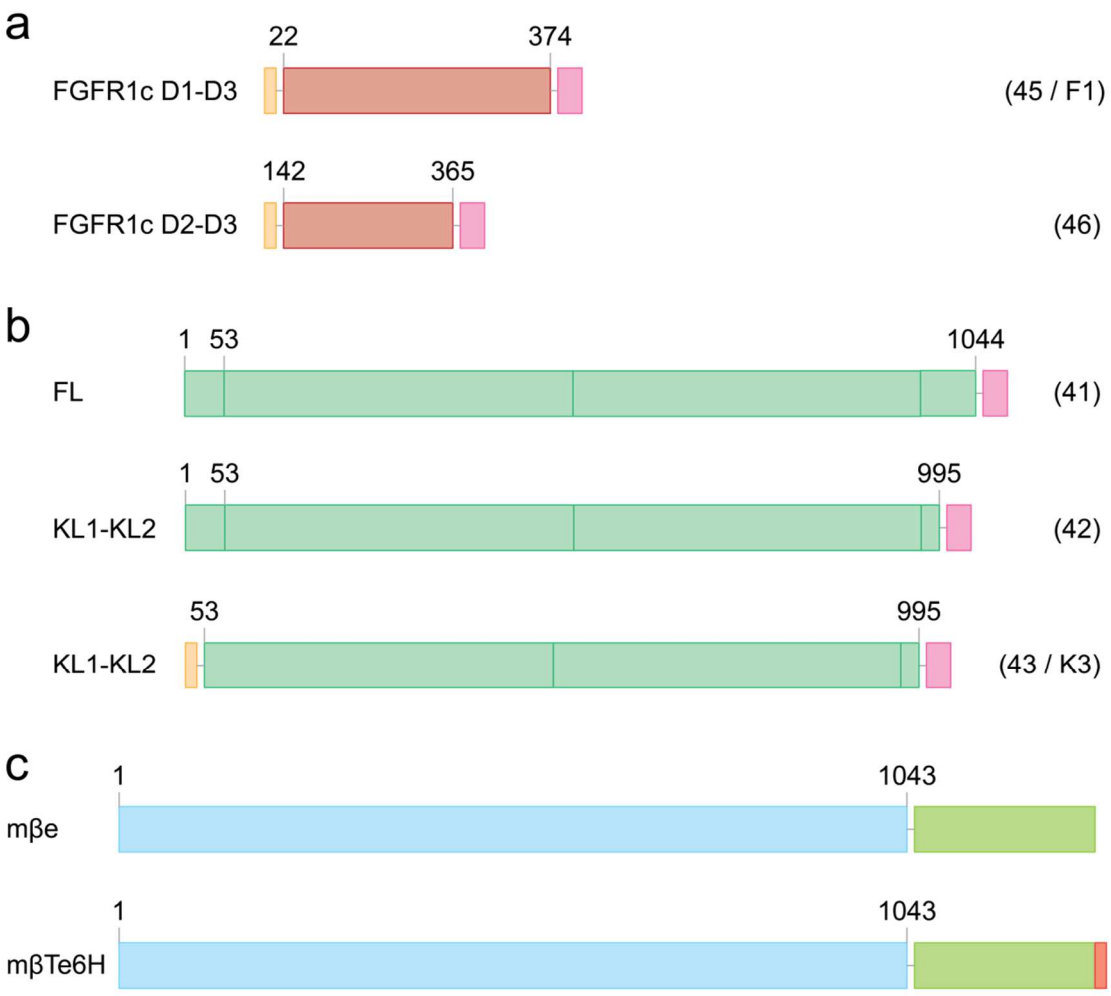


Figure 4.18 Mammalian expression constructs: design and nomenclature. Fundamental design of FGFR1c and βKlotho constructs designed for mammalian expression. All human constructs (**a**-**b**) were codon optimised for expression in mammalian cells and comprised a TEV protease-cleavable C-terminal 2StrepII-10His tag (pink). Non-membrane bound extracellular constructs were also generated with the native (βKlotho, green) or a non-native (human CD33, orange) N-terminal secretion signal peptide. **(a)** Human FGFR1c extracellular domain constructs boundaries were the same as those used for bacterial expression. **(b)** Human βKlotho construct boundaries were the same as those used for the bacterial expression but with a focus on the longer KL1-KL2 constructs and the full length protein. **(c)** A full length mouse βKlotho cDNA construct with a C-terminal fusion to eGFP (mβe) was obtained (Addgene). A TEV-protease cleavable C-terminal 6His tag (bright red) was added to aid purification (mβTe6H). For full experimental details, see section 2.2.2 and refer to Table 2.4 for further construct information.

These constructs were then put through a series of small scale (10 mL) expression tests, by transiently expressing the proteins first in FreeStyle™ 293F cells and later in Expi293F™ and CHO-EBNA GS cells. Samples of whole cells and media were analysed for each expression. Samples were taken every day for 3-4 days post-transfection for FreeStyle™ 293F expression tests. Samples were taken at 3 days (for proteins expected to be intracellularly expressed) and 7 days (for proteins expected to be secreted into the

medium) post-transfection for Expi293F™ and CHO-EBNA GS expression tests. Representative results for all constructs with observable expression are summarised in Figure 4.19. In FreeStyle™ 293F cells, the longer length FGFR1c D1-D3 protein was reproducibly secreted in large quantities into the medium, while the shorter length FGFR1c D2-D3 protein was also secreted into the medium but in lower quantities and less reproducibly (Figure 4.19a). In Expi293F™ and CHO-EBNA GS cells a similar pattern of secretion and reproducibility was observed for these two proteins, one of which could be expressed in both cell types and from two different expression vector backbones (Figure 4.19b). Disappointingly, none of the human βKlotho proteins could be detected in either cells or media in any expression tests performed in FreeStyle™ 293F cells (data not shown). Very little mβTe6H expression could be detected in these cells (Figure 4.19d). Some of the human and mouse βKlotho proteins were detected in expression tests in Expi293F™ and CHO-EBNA GS cells, however those expected to be secreted into the medium were detected intracellularly (Figure 4.19d and 4.19f). Additionally, an experiment performed in Expi293F™ cells co-transfected with the FGFR1c D1-D3 construct and one of the βKlotho KL1-KL2 constructs revealed that while co-expression appeared to reduce the production of secreted FGFR1c, larger quantities of βKlotho KL1-KL2 could be detected in the co-transfected cells, albeit still not secreted into the medium as should be expected (Figure 4.19e). Preliminary attempts were made to extract and purify all these visibly expressing extracellular KL1-KL2 and FL βKlotho proteins from cell lysates but this proved to be challenging and none of the proteins could be extracted into the soluble fraction from these samples (Figure 4.19g). Purification of the abundant extracellular FGFR1c proteins was successful, and is detailed in section 4.3.5.

Little expression of the full length human and mouse βKlotho proteins could also be detected in both Expi293F™ and CHO-EBNA GS cells (Figure 4.19f). However, it was not possible to extract these proteins into the soluble fraction for purification using standard extraction methods (Figure 4.19g). It was hypothesised that suspension cells might not provide an optimal environment to express these membrane proteins. To determine whether this might be the case, expression of both the mβe and the mβTe6H constructs was tested in HEK 293T cells and followed by monitoring the green epifluorescence of the C-terminal GFP fusion tag. In these cells, both mouse βKlotho constructs were expressed and the proteins appeared to localise to the plasma membrane as expected (Figure 4.20a-b). This promising result led to a medium-scale expression trial of mβTe6H in these cells, followed by a detergent extraction of the membrane protein using a Triton X-100 based chemical lysis similar to that used for all

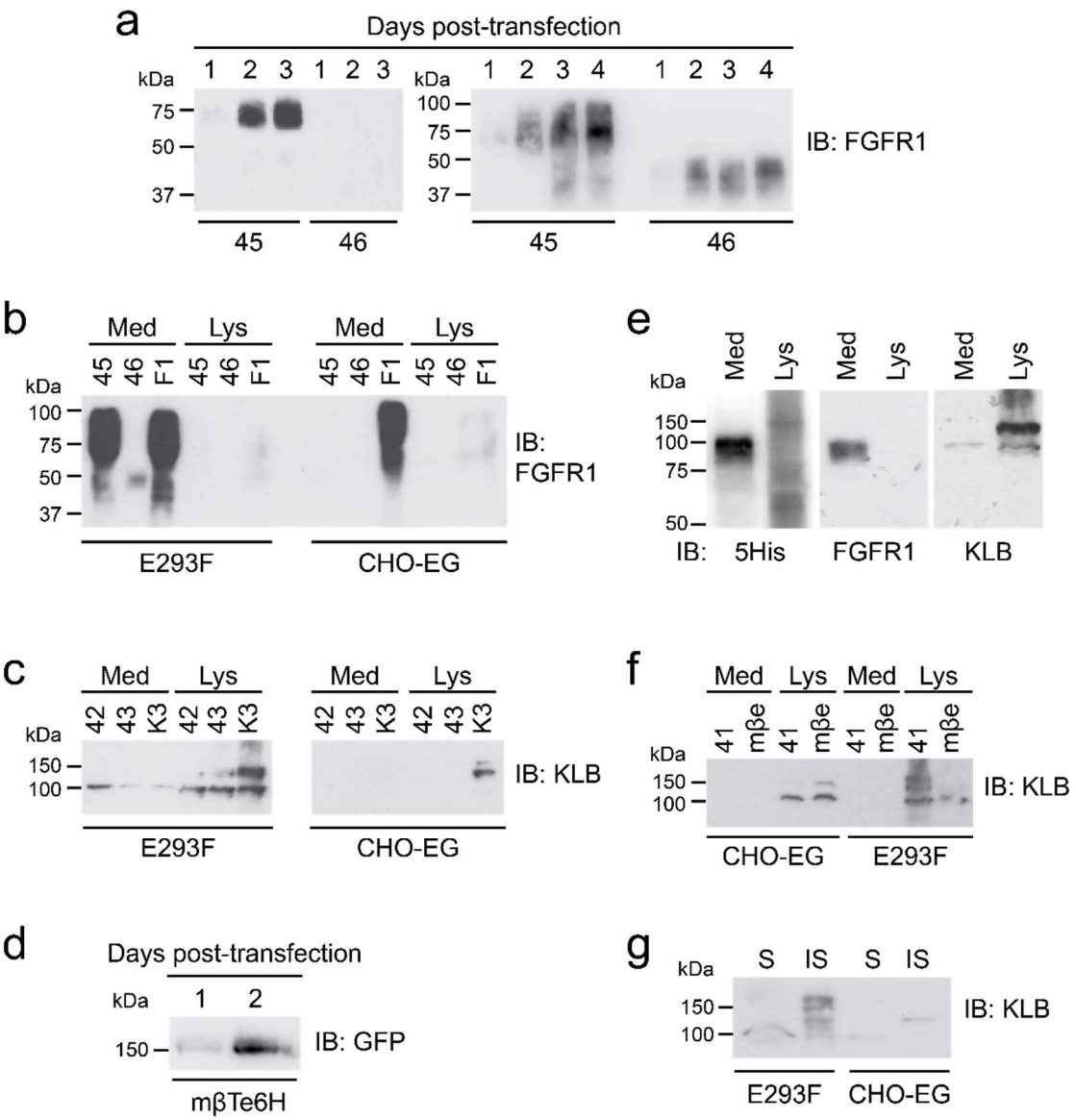

a
Days post-transfection
kDa
1 2 3 1 2 3
75
50
37
45
46
kDa
1 2 3 4 1 2 3 4
100
75
50
37
45
46
IB: FGFR1
b
Med
Lys
Med
Lys
45 46 F1 45 46 F1 45 46 F1 45 46 F1
kDa
100
75
50
37
IB: FGFR1
E293F
CHO-EG
e
Med
Lys
Med
Lys
Med
Lys
kDa
150
100
75
50
IB: 5His
FGFR1
KLB
c
Med
Lys
Med
Lys
42 43 K3 42 43 K3 42 43 K3 42 43 K3
kDa
150
100
IB: KLB
E293F
CHO-EG
f
Med
Lys
Med
Lys
41 mβe 41 mβe 41 mβe 41 mβe
kDa
150
100
IB: KLB
CHO-EG
E293F
d
Days post-transfection
kDa
1 2
150
IB: GFP
mβTe6H
g
kDa
S IS S IS
150
100
IB: KLB
E293F
CHO-EG

Figure 4.19 Successful recombinant protein expression from mammalian expression tests in HEK 293F and CHO cells. From the panel of mammalian expression constructs generated (Table 2.4 and Figure 4.18) those that expressed from FreeStyle™ 293F (F293F), Expi293F™ (E293F) or CHO-EBNA GS (CHO-EG) suspension cells are depicted here by immunoblotting using the indicated antibodies. Cells were transiently transfected on Day 0 and samples were taken 1-7 days post-transfection. Transfection of E293F and CHO-EG cells was performed at AstraZeneca using their cell lines under the supervision of David Fisher. Secreted protein expression was analysed in conditioned media (Med) while intracellular protein expression was analysed in whole cell lysates (Lys) comprising both soluble (S) and insoluble (IS) fractions. Numbering of constructs is as in Table 2.4 and Figure 4.18. In all anti-KLB immunoblots, the lower molecular weight band detected in E293F samples (at approximately 100 kDa) corresponds to a non-specific band (validated by probing a non-specific lysate with the anti-KLB antibody, data not shown). **(a)** Secreted expression for FGFR1c constructs from F293F cells with samples taken 1-4 days post-transfection. Immunoblots show results from two out of three independent expression tests for each construct, demonstrating the reproducibility of expression of the D1-D3 construct (45) while the D2-D3 construct (46) only expressed in two out of three expression tests. **(b)** Expression of all FGFR1c constructs from E293F and CHO-EG cells with samples taken 7 days post-transfection. **(c)** Expression of all human βKlotho extracellular constructs from E293F and CHO-EG cells with samples taken 7 days post-transfection. **(d)** Expression of mouse full length βKlotho constructs from F293F cells with samples taken 1-2 days post-transfection. **(e)** Co-expression of human FGFR1c and βKlotho constructs F1 and K3 (this transfection was performed at AstraZeneca by David Fisher) from E293F cells with samples taken 3 days post-transfection. **(f)** Expression of human and mouse full length βKlotho constructs from E293F and CHO-EG cells with samples taken 3 days post-transfection. **(g)** The mβe, 41 and K3 Lys samples from E293F and CHO-EG cells were further separated into S and IS fractions to test extraction and purification for each protein. The non-specific band was found in the S fraction while the Klotho bands were found in the IS fraction. The immunoblots show results for mβe, representative for S/IS separations of all three proteins. For full experimental details, see section 2.17 and refer to Table 2.4 for construct information.

other proteins in this work. The protein was only partially extracted into the soluble fraction by this method and the resulting solubilised protein proved challenging to purify further (Figure 4.20b, right panel). To further optimise the purification procedure, a detergent screen was performed to determine if the protein could be better extracted using a different detergent (Figure 4.20c). Promising preliminary data emerged from this screen, as all the detergents screened performed better at mβTe6H extraction than Triton X-100, with the non-ionic detergent *n*-decyl-β-D-maltopyranoside (DM) providing the best extraction. Further experiments using these detergents, and possibly their combinations, to extract βKlotho proteins might eventually lead to success in purifying this challenging protein from HEK 293 cells.

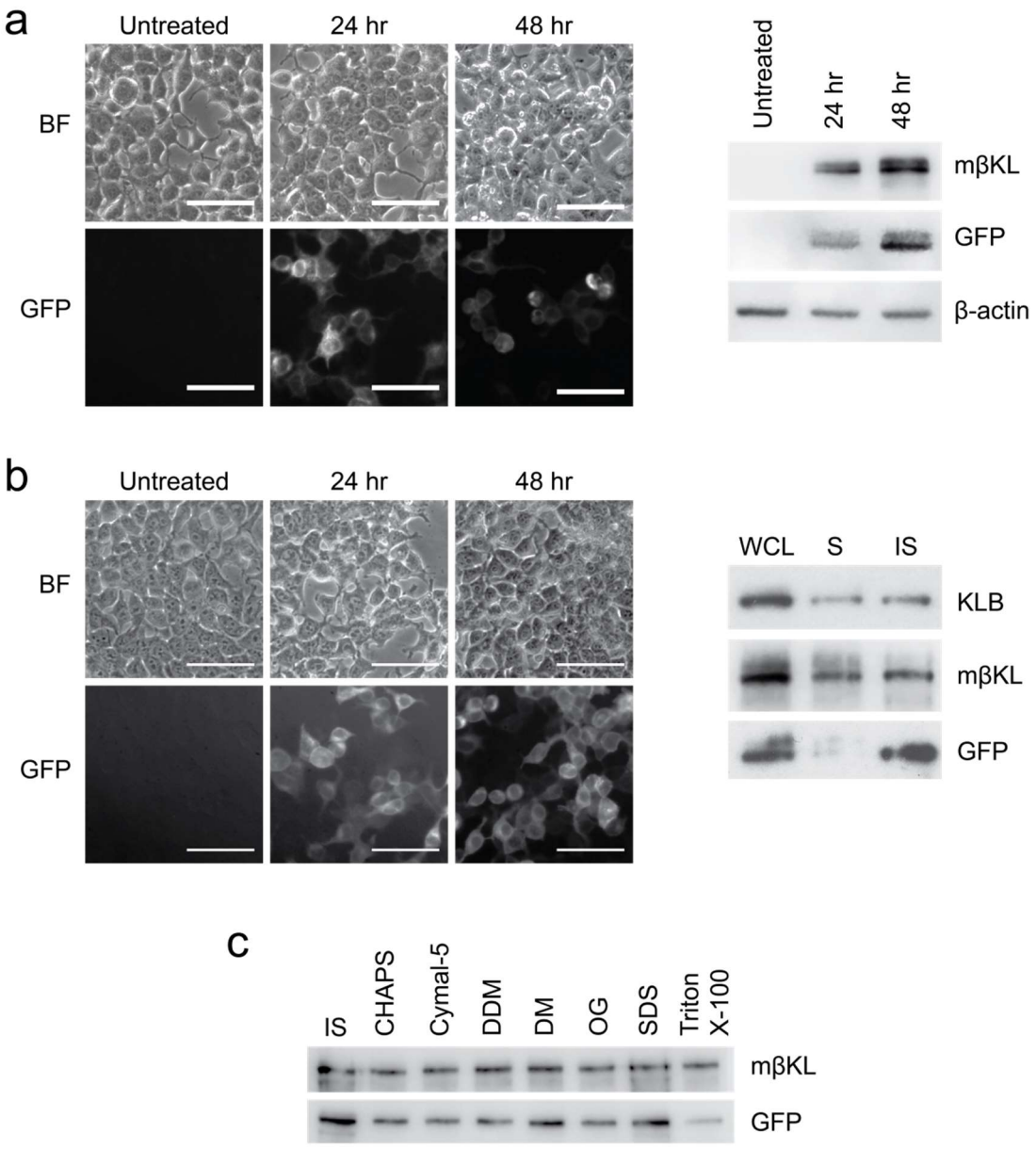


Figure 4.20 Extraction of GFP-tagged mouse βKlotho expressed in HEK 293T cells. HEK 293T cells were transiently transfected with each GFP-tagged mouse βKlotho construct and expression was monitored visually by green epifluorescence microscopy and biochemically by immunoblotting using the indicated antibodies over a 48 hour time period. Scale bars represent 62.5 µm. BF – brightfield, GFP – green fluorescence protein (green epifluorescence). **(a)** Cells transiently transfected with the mβe construct (or a non-transfected control) for 24 or 48 hours showed mβe expression through green epifluorescence (left panel) and immunoblotting of whole cell lysates (WCL) with the indicated antibodies (right panel). **(b)** Cells transiently transfected with the mβTe6H construct (or a non-transfected control) for 24 or 48 hours showed mβTe6H expression through green epifluorescence (left panel) and immunoblotting of WCL, S and IS fractions with the indicated antibodies (right panel). Results are representative of four independent experiments. **(c)** Detergent screen for solubilisation of mβTe6H expressed in the insoluble (IS) fraction. Solubilised mβTe6H was detected by immunoblotting using the indicated antibodies. CHAPS – (3-((3-cholamidopropyl)-dimethylammonio)-1-propanesulfonate), Cymal-5 – 5-cyclohexyl-1-pentyl-β-D-maltoside, DM – *n*-decyl-β-D-maltopyranoside, DDM – *n*-dodecyl-β-D-maltoside, OG – octyl glucoside, SDS – sodium dodecyl sulphate. For full experimental details, see sections 2.10 and 2.14.

4.3.5 Purification of FGFR1c proteins

As demonstrated in section 4.3.4, extracellular FGFR1c proteins can be recombinantly expressed in suspension mammalian cells (Figure 4.19a-b). The longer length FGFR1c D1-D3 protein was abundantly secreted from all three cell types tested, whereas the shorter length D2-D3 protein was less robustly expressed. Attempts to purify the D2-D3 protein proved challenging and due to the lower yield of this protein further purification was not attempted from the small-scale expression tests (data not shown). However, the more-physiologically relevant FGFR1c D1-D3 proteins were all successfully purified from all expression test media samples using a combination of affinity chromatography and SEC (detailed methods in section 2.18). SEC chromatograms, Coomassie stained SDS-PAGE gels and immunoblots used to demonstrate protein purity are shown in Figure 4.21. Protein purified from CHO compared to 293F cells ran at a slightly higher molecular weight on the gels, possibly revealing different glycosylation states of protein produced by the two cell types. Expression in the Expi293F™ cell line yielded the largest quantities of FGFR1c protein at 60 and 130 mg/L (respectively from pcDNA3.1 and pEBNAZ expression vector backbones), while the FreeStyle™ 293F cells yielded 30 mg/L and CHO-EBNA GS cells yielded 45 mg/L. The sheer abundance of these proteins facilitated their successful purification from just 10 mL expression cultures. All four FGFR1c D1-D3 protein preparations were estimated from SDS-PAGE analysis to be more than 95% pure (Figure 4.21b) and suitable for use in further experiments.

4.3.6 Preliminary characterisations of endocrine FGF-FGFR1c binary complexes

To begin to understand the protein-protein binding interactions involved in endocrine FGFR signalling complexes, initial characterisations of the endocrine FGF-FGFR1c binary complexes were performed using biophysical methods. As it was not possible to purify any βKlotho proteins to completion, no *in vitro* binding experiments could be performed in this work to characterise the FGF-FGFR1c-βKlotho ternary signalling complexes. The binary heteromeric complexes formed between FGF19-FGFR1c and FGF21-FGFR1c, prepared with recombinant proteins purified in this work, were analysed to begin to elucidate their binding affinities and thermodynamic and kinetic properties.

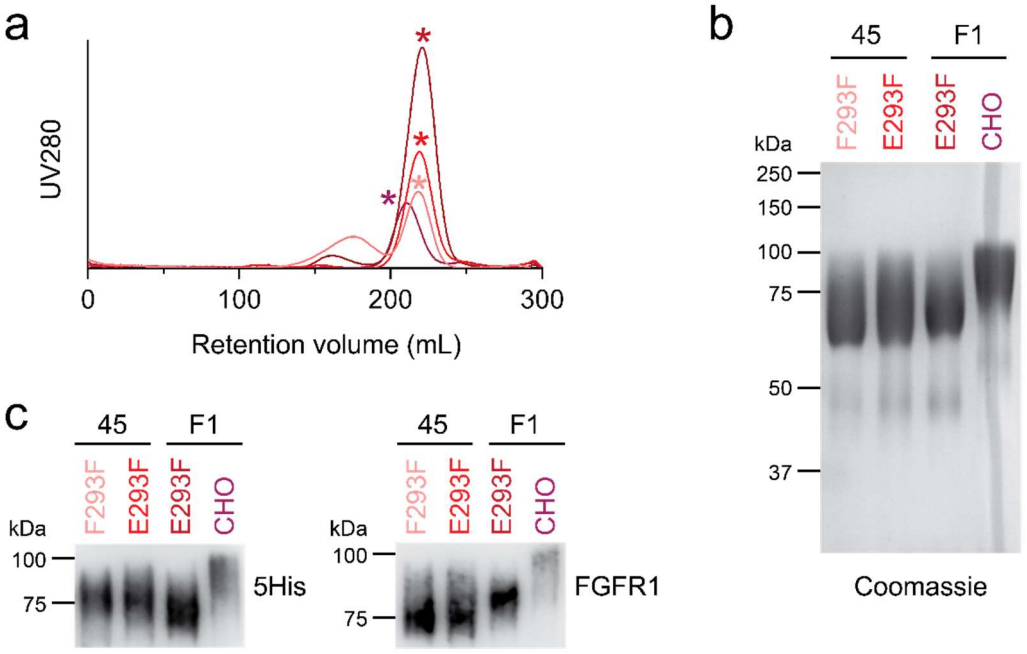


Figure 4.21 Purification of FGFR1c proteins. (a) Overlay of size exclusion chromatography (SEC chromatograms for the final stage of purification of all four recombinant FGFR1c D1-D3 proteins expressed in mammalian suspension cell lines. A single asterisk highlights the peak containing monomeric, soluble FGFR1c protein for each preparation. **(b)** Purified FGFR1c proteins from the SEC peaks marked by a single asterisk in **a** from expression in all four mammalian cell lines visualised by Coomassie stained SDS-PAGE analysis. From this analysis, protein was estimated to be more than 95% pure. **(c)** Immunoblotting using the indicated antibodies on the same purified protein samples as in **b**. For full experimental details see section 2.18.

One of the best methods to study such biophysical properties is isothermal titration calorimetry (ITC). This technique can report on all parameters of binding interactions concurrently and in a single experiment, such as affinity, enthalpy, entropy and stoichiometry of binding of biomolecules in their native label-free state in solution. ITC measures the heat transfer during binding by gradually titrating one binding partner into the other in solution, providing a thermodynamic profile of the binding interaction. For these reasons, ITC was employed as a starting point for quantitatively characterising these protein-protein interactions. The MicroCal™ iTC-200 instrument (Malvern) was used for these experiments. FGF1 and FGF21 ligands were independently titrated into FGFR1c in the sample cell in two independent experiments each. Unfortunately, it was not possible to measure heats of interaction for either of the FGF ligands in the single buffer and at the single temperature (25°C) tested, despite a contrived fit attempted for FGF1 (Figure 4.22). It was thought that optimising these experiments further might consume too much protein sample, of which the FGFs were particularly challenging to prepare. Therefore, a different approach was considered instead.

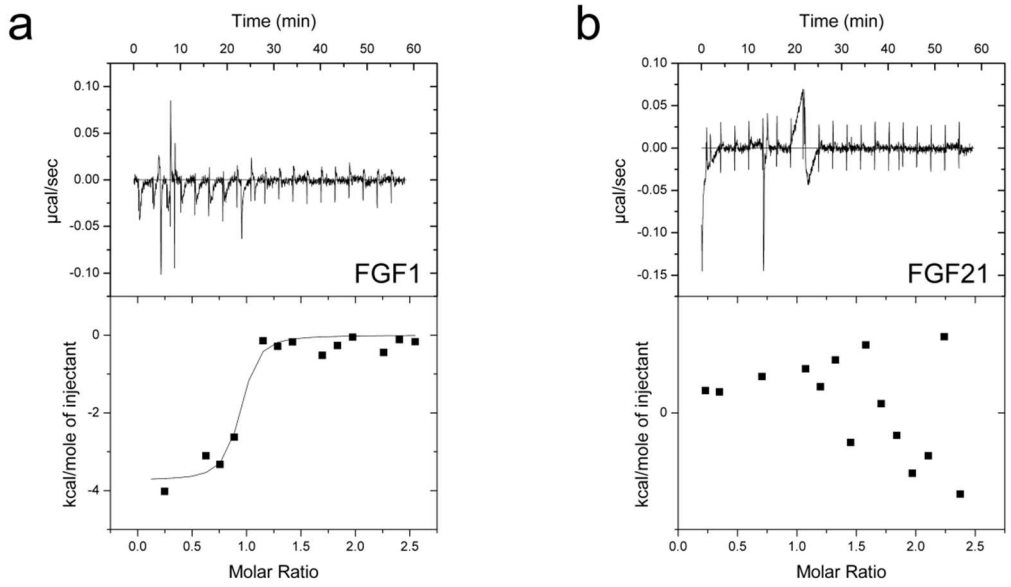


Figure 4.22 Isothermal titration calorimetry (ITC) experiments were unsuccessful for the interrogation of FGF-FGFR binary complex interactions. Pilot ITC experiments performed using FGFR1c and a ten-fold molar excess of recombinant FGF ligands **(a)** FGF1 or **(b)** FGF21 were largely unsuccessful for probing the binding interactions between these proteins using the single reaction buffer and temperature tested. For full experimental details, see section 2.20.3.

Previous attempts have been made to characterise FGF-FGFR1c-βKlotho interactions using surface plasmon resonance (SPR) experiments (Goetz et al., 2012b), albeit using the shorter FGFR1c D2-D3 protein rather than the longer and more physiologically relevant FGFR1c D1-D3 protein prepared in this work. In order to determine whether it was possible to reproduce these data, preliminary SPR experiments were performed with FGFR1c D1-D3, FGF19 and the two independent preparations of FGF21 prepared in this work. The Biacore™ T200 system (GE Healthcare) was used for the experiments. Each FGF was immobilised (using the C-terminal 10His tag) on three out of four flow cells of an NTA sensor chip, leaving the first flow cell empty as a reference cell to record background response including non-specific binding of FGFR1c D1-D3 (the analyte) to the chip surface. Different concentrations of analyte were then flowed simultaneously over each flow cell using a flow rate of 30 µL/min. Representative sensorgrams and equilibrium dissociation constant (K_D) values for each interaction are shown in Figure 4.23. The maximum response (R_{max}) for all active flow cells was lower than expected based on the levels of FGFs immobilised (Table 4.6) suggesting that the FGFs were only partially or heterogeneously active. A decrease in response was observed at the end of each analyte injection period (particularly at higher analyte concentrations) which is likely

due to analyte protein sticking to the microfluidics. To minimise inaccuracy, the steady state response values used to calculate K_D values were taken at 90 seconds post-injection of analyte. Although it was possible to determine preliminary K_D values for the FGF-FGFR interactions from these pilot data, it was not possible to fit the data to any kinetic model within the Biacore™ T200 evaluation software. This might be due to FGF ligand heterogeneity or possibly due to transient homodimerisation of FGF and/or FGFR1c analyte proteins in addition to FGF-FGFR heterodimerisation. The K_D values measured here were two orders of magnitude lower than those determined previously (Figure 4.23 and Table 4.6), although the FGFR1c construct used to assess binding in this work is longer by an entire Ig-like domain than the FGFR1c protein used in the previous study which lacks the D1 domain (Goetz et al., 2012b). Although early evidence proposed roles for the D2 and D3 domains in binding to the endocrine FGFs, this was proposed following experiments performed using FGFR proteins lacking the D1 domain, and therefore the potential importance of the D1 domain in binding to endocrine FGFs cannot be ruled out. Perhaps the preliminary K_D values measured here (Figure 4.23 and Table 4.6) do indeed suggest a putative role for the FGFR D1 domain in conjugation of FGFs within the endocrine ternary complexes, but need to be directly compared with K_D values measured with the D2-D3 protein as well to confirm this. These pilot data provide a useful starting point to further probe these interactions. Further optimisations of these experiments will provide more accurate information about the true binding affinities and kinetics involved in binding of the endocrine FGFs to FGFR1c. This will help to begin characterisation of the protein-protein interactions involved in the binary and ternary FGFR endocrine signalling complexes.

Table 4.6 Comparison between preliminary equilibrium dissociation constants (K_D) for endocrine FGF-FGFR binary interactions measured by surface plasmon resonance (SPR) in this work (± standard error) and by a similar study. Other SPR parameters from the preliminary experiments in this work are also reported. RU – response units.

	Goetz *et al.*, 2012 hFGFR1c (142-365)	This work hFGFR1c (22-374)	Immobilised FGFs (RU)	R_{max} (RU)
K_D FGF19-hFGFR1c	1.91 µM	68.1 ± 2.1 nM	12.6	6.4
K_D FGF21-hFGFR1c	2.52 µM	53.8 ± 2.2 nM	13.9	13.7
		44.0 ± 2.1 nM	9.4	12.4

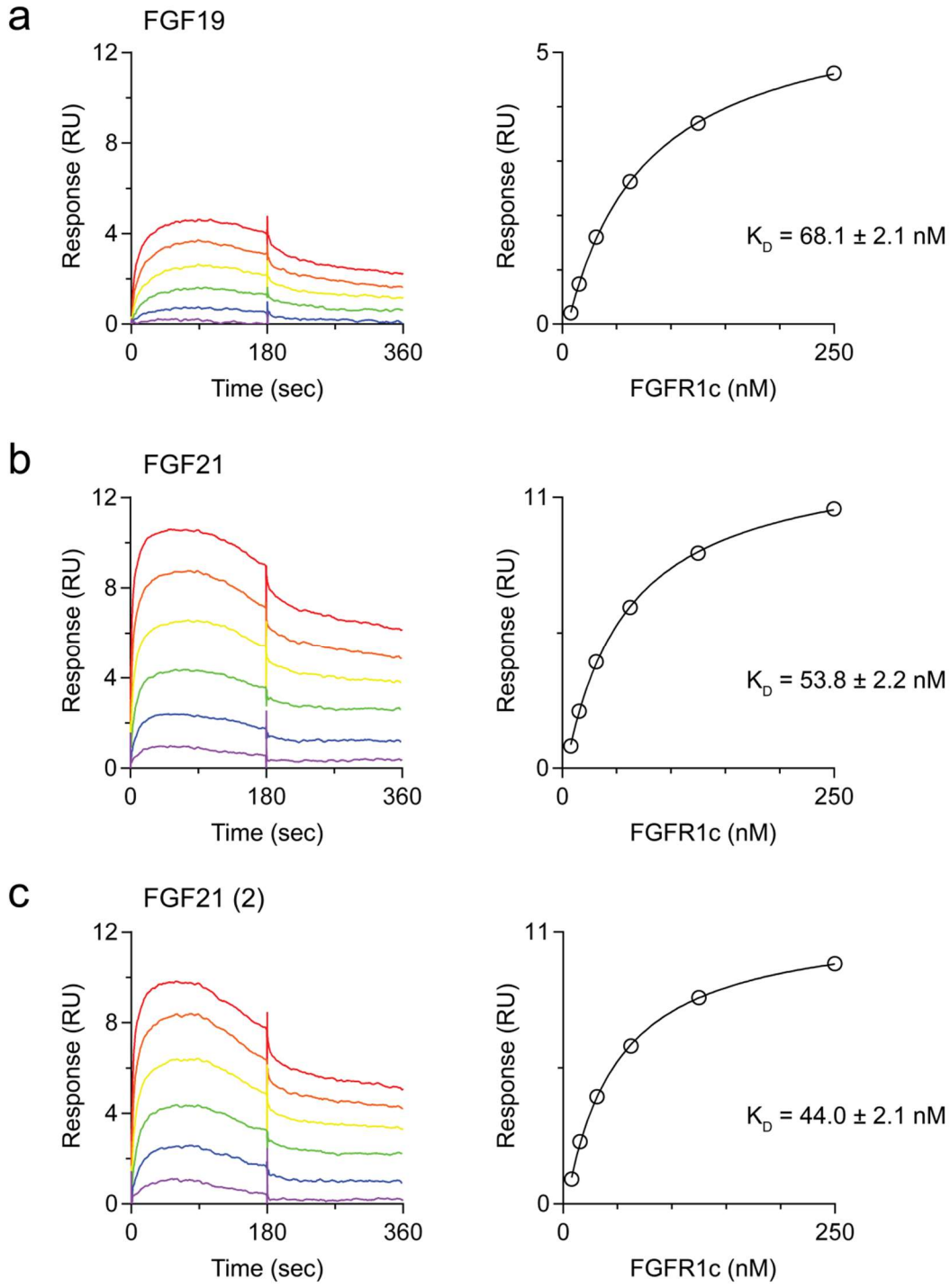


Figure 4.23 Surface plasmon resonance (SPR) experiments provide preliminary binding affinities for the endocrine FGF-FGFR binary interactions. Preliminary SPR experiments were performed by immobilising recombinant **(a)** FGF19, **(b)** FGF21 and **(c)** FGF21 (2) on three flow cells of an NTA sensor chip (Table 4.6) using their C-terminal 10His tags, leaving the first flow cell empty to serve as a reference flow cell. Recombinant FGFR1c (from F293F cells, Figure 4.21) at increasing concentrations between 0-250 nM (purple – red traces) was then flowed over the flow cells and responses recorded. Sensorgrams (left panels) and steady-state responses used to calculate the equilibrium dissociation constants (K_D) (right panels) are shown and are representative of two independent experiments performed under slightly different buffer conditions (0.05% and 0.005% Tween® 20). The K_D values ± standard error calculated from the non-linear fits are indicated. Data were analysed using the Biacore™ T200 software (Biacore). For full experimental details, see section 2.20.4.

4.4 Discussion

A wealth of research into the canonical modes of paracrine FGFR signalling has shed light on its importance in all aspects of developmental and metabolic homeostasis. However, the endocrine sub-axes of FGFR signalling have been much less well researched, particularly from a biochemical and structural perspective. Therefore, despite the importance of these signalling complexes in regulating a number of vital metabolic processes and malignant outcomes, a lot of molecular-level information linking structure to function is currently lacking. This information is required to elucidate the molecular mechanisms by which the Klotho co-receptor family play a role in both normal and deregulated FGFR signalling. This will inform the design of endocrine FGF biologic drug candidates and anti-endocrine FGF therapeutics currently in pre-clinical and clinical development for the treatment of metabolic diseases and cancer.

In this chapter the foundations have been laid to investigate endocrine FGFR signalling complexes from a structural, biophysical and biochemical perspective. There was a focus on investigating the sub-axes of signalling complexes involving the most important FGFR, FGFR1c, with its co-receptor βKlotho and the cognate ligands to their binary complexes, FGF19 and FGF21. Combinations of bioinformatics and literature searching methods were used to determine construct lengths for the production of expression competent constructs for all three protein components of the ternary complexes. *In vitro* experiments were employed to begin to establish analytical routes to study these protein complexes further. The experimental part of this work proved to be very challenging, involving a lot of method development as well as optimisation of previously published protocols which often omit the detailed experimental methods needed to reproduce them. This work progressed from the difficult molecular biology stages to preliminary biophysical characterisations of binary complexes with two out of three purified components of the ternary signalling complexes. Robust methods to express and purify these two components have been presented, and progress was made in recombinantly producing human βKlotho proteins from *E. coli* and mammalian cells. The most encouraging results for production of the larger extracellular proteins came from expression tests in mammalian cells, which have been less commonly used in the literature to produce these recombinant proteins for *in vitro* studies. Several methods and experiments that proved unsuccessful with these materials can be ruled out for future research involving these complexes. In this section, both the technical and theoretical implications of this work will be discussed in more detail.

4.4.1 Bioinformatic analyses used to guide experimentation and drug discovery

There is a complete lack of structural information available for all Klotho proteins and any of the individual KL domains. This could be because, as in this work, the Klotho KL domains and longer Klotho proteins proved challenging to produce recombinantly in terms of cDNA manipulation, molecular biology and stages of protein overexpression and purification. Structural information of the extracellular Klotho domains would aid in the rational design of endocrine FGF related therapeutics for the treatment of several diseases including diabetes, obesity, cancers and genetic diseases (see section 4.1). To alleviate some of the structural challenges, a combination of sequence analyses, *in silico* modelling approaches and comprehensive literature searching were undertaken for βKlotho and its individual KL domains in order to inform experimental planning. Multiple sequence alignments, secondary structure prediction tools and homology modelling based on the most closely related protein structures available in the RCSB PDB provided a useful starting point to determine probable KL domain boundaries and therefore protein construct boundaries for recombinant expression. The low degree of sequence conservation between KL1 and KL2 domains, despite a high degree of apparent structural similarity, made boundary selection challenging. Evaluation of βKlotho proteins expressed and purified by others (Table 4.4) coupled with multiple bioinformatics analyses proved most useful in helping to define the lengths of KL domain proteins expressed in this work. Further fine-tuning of the construct boundaries might help to improve soluble yields of the individual KL domain proteins; this remains to be tested experimentally. Nevertheless, the fact that some degree of expression and purification was possible for both KL1 and KL2 domain proteins might suggest that more stringent chromatographic purification will eventually lead to the production of purer protein preparations. This might also facilitate resolution of the protein aggregation problem or alternatively, detergent screening during protein extraction (similar to that performed for the full length mouse βKlotho) might help identify different routes to purify these proteins to a final soluble and non-aggregated state.

The homology models of the βKlotho KL1 and KL2 domains generated in this work provided valuable information about expected secondary and tertiary structure of these experimentally structurally uncharacterised domains. This was useful in developing protein purification methods and will complement future experimental structural studies. The models can be used to aid the design of site-directed mutagenesis experiments to elucidate KL domain function and probe inter-protein binding interfaces. The models are also very similar to models created recently for the αKlotho KL1 domain (Wright et al., 2017) and part of the βKlotho KL1 domain (Ming et al., 2012) adding a level of validity in

using these models to inform structure-based design of βKlotho related therapeutics in a case where atomic structures might be challenging to obtain.

Sequence divergence of the two βKlotho KL domains might suggest that each KL domain has sequence specific roles, for example in ligand binding or dimerisation. Therefore although the primary aims of this work were to reconstitute the entire extracellular (or full length) endocrine signalling complexes, it was interesting to additionally prepare proteins of each individual KL1 and KL2 domain with the aims of elucidating their potential roles. Although optimisations are still required to purify soluble monomeric versions of these recombinant proteins (discussed above), the basis for expression and purification from *E. coli* has been established in this work. Additionally, secretion of these KL domains from mammalian cell systems should be considered to produce sufficient protein as required for structural and biophysical experiments. This was done recently to recombinantly produce a secreted splice variant of the αKlotho KL1 domain (Wright et al., 2017), although the authors gave no indication of the final protein yields from expression in HEK 293 cells. With sufficient protein, further work might then try to define the roles of each KL domain in binding to FGFs, FGFRs and possibly in βKlotho homodimerisation, if this occurs (currently unknown). Paired with *in silico* protein-protein docking, site-directed mutagenesis and solvent exclusion studies (such as using hydrogen-deuterium mass spectrometry or nuclear magnetic resonance experiments), it should soon be possible to define the specific sequences involved in the interactions of the ternary complexes. Simple native PAGE or analytical SEC experiments could also be used to assess the ability of each KL domain to dimerise with other KL domains and/or FGFs and/or FGFRs.

Another debate that remains unresolved concerns the implied β-glucuronidase and α2,6-sialidase enzymatic activities of the KL domains (Chang et al., 2005, Tohyama et al., 2004, Cha et al., 2008) despite these domains lacking crucial catalytic aspartates needed for glycosyl hydrolase activity (Figure 4.2). A recent study suggested that a secreted version of the αKlotho KL1 domain may be sufficient for binding to α2-3-sialyllactose (Wright et al., 2017). This adds to the range of substrates implicated in being bound and/or metabolised by Klotho proteins. It might also exemplify the possibility of sequence specific roles for the different KL domains. Indeed, KL domains from different Klotho family members might target diverse and context-dependent substrates depending on, for example, tissue specific expression. An example from within the β-glucosidase type I family comes from the *S. alba* myrosinase. This enzyme is a β-glucosinolase but one of the two essential glutamate residues conserved across the rest of the β-glucosidase type I family has been replaced by glutamine in the myrosinase (Burmeister et al., 1997).

Instead, this enzyme has a modified function as an *S*-glycosidase (instead of an *O*-glycosidase) and catalyses its substrate with the help of an ascorbate co-factor that replaces its need for the 'missing' catalytic aspartate (Burmeister et al., 2000). Similarly for the KL domains, sequence variation of the catalytic aspartates may indeed have altered putative enzymatic activities of the KL domains. Alternatively, combined with the existence of multiple putative glycosylation sites on the Klotho proteins (Figure 4.2 and Table 4.1), the deviation of catalytic residues from consensus sequences could have important implications for cell surface signalling whereby binding to various carbohydrate moieties might play vital roles in determining the tissue specificity of Klotho related signalling, without the necessity for carbohydrate catabolic activities. High-throughput screening of carbohydrates for their ability to bind Klotho proteins or individual KL domains, coupled with *in silico* docking studies using different carbohydrate moieties and the homology models produced in this work, should help to narrow down possible candidates for further testing in quantitative *in vitro* enzyme-substrate metabolic activity assays. This, combined with future structural information, should help to resolve this currently open question.

4.4.2 Refolding inclusion bodies: an empirical method of 'last resort'?

Attempts were made to refold both endocrine FGFs and FGFR1c proteins expressed in inclusion bodies in *E. coli* in this work. Despite closely following what was available of published protocols (see section 4.3.2), it proved to be challenging to produce sufficient amounts of folded FGFR1c protein via such methods. Refolding proteins from inclusion bodies might be an empirical method of 'last-resort' which can be employed successfully for proteins such as relatively small growth factors and even FGFs, as reported above, but has proved difficult with larger proteins. In this work, the problems incurred with trying to refold FGFR1c proteins led to attempts to express them using secretion from mammalian expression systems. All suspension mammalian expression systems tested in this work yielded large quantities of soluble, non-aggregated FGFR1c protein. In fact, superior expression was observed for the longer length extracellular FGFR1c protein via these methods in terms of both yield and quality of the protein; this is the more physiologically relevant construct compared to the shorter length D2-D3 protein used in the majority of previous structural studies and binding experiments in the literature (section 4.3.2). Mammalian systems are rapidly gaining popularity for protein overexpression, even for structural studies, perhaps because the relative length of time and costs involved compared to bacterial expression systems might be worthwhile if the ultimate rates of success in robust protein production are higher. These potential benefits remain to be evaluated including for the proteins produced in this work. Additionally, for biochemical and biophysical experiments using human proteins, recombinant proteins

produced in mammalian cells are more likely to be natively or near-natively post-translationally modified, including glycosylation or smaller chemical modifications such as phosphorylation or acetylation. Modifications such as glycosylation can be crucial for protein function and intra- and inter-complex interactions of the ternary signalling complexes being investigated in this work, as both FGFRs and Klotho proteins have several glycosylation sites on their extracellular domains (see Figure 4.2 and Table 4.1 for Klotho glycosylation information) with mostly uncharacterised specific functions. Glycosylations might be crucial players in mediating tissue-specific functions of FGFR endocrine signalling complexes, including in metabolic tissues and in specific cancers.

4.4.3 Evaluation of quality control of refolded proteins

When refolding proteins from denatured inclusion bodies, it is vital to demonstrate that the final purified protein is correctly folded, homogeneous, and bioactive. Successfully demonstrating bioactivity would imply that protein is also correctly folded. Therefore cell-based assays were developed to test the bioactivity and indirectly the 'foldedness' of the FGFs produced in this work. The 3T3-L1 adipocytes used were cultured and differentiated using an established protocol with slight modifications (Zebisch et al., 2012) and all experimental parameters were similar to those used previously (Kurosu et al., 2007, Ogawa et al., 2007, Yang et al., 2012). Despite successful differentiation of the pre-adipocytes to adipocytes as visually confirmed by phase contrast microscopy (data not available), problems were encountered with both bacterial and fungal infections in this cell type, which worsened throughout the differentiation process. Although only cells with no visible infection were used in these experiments, underlying infections could not be ruled out, and might explain the unexpected responses of the cells to the different FGFs including lack of dose response with high concentrations of FGF1 and a general lack of response with commercially obtained control endocrine FGFs (Figure 4.15). In this case, being able to assess secondary structural fold of the recombinant FGFs using biophysical methods was very useful in validation of the refolding methods used here, despite only providing little and indirect information about bioactivity of the refolded FGFs. The SPR experiments revealed that the endocrine FGFs were probably only partially active or conformationally heterogeneous (section 4.3.6). This might reflect the intrinsic problem with *in vitro* protein refolding leading to the creation of a heterogeneously folded mixture of proteins. The homogeneity of this mixture could be improved by adding orthogonal chromatographic techniques to the purification such as ion exchange and reverse phase chromatography, although these were unsuccessful in this study and dramatically reduced the final purified protein yields (data not shown). Alternatively, it might reflect the fact that both refolded FGFs demonstrated instability as they visibly degraded over time with FGF19 degrading faster than FGF21 (Figures 4.13

and 4.14), and in fact this was also observed for recombinant FGF19 and FGF21 proteins produced by others, for example (Asada et al., 2009). Protein degradation over time might therefore be a more plausible explanation since all FGFs appeared to be correctly folded when assessed by CD spectroscopy (Figure 4.17) but then appeared to be heterogeneous during SPR experiments performed a few weeks later. Interestingly, other groups have used mammalian, insect and yeast cell systems to express recombinant endocrine FGFs (Xie et al., 1999, Inagaki et al., 2005, Micanovic et al., 2009, Wu et al., 2011). It could be worth pursuing this experimental course considering the success of mammalian expression systems in this work in producing large quantities of the extracellular FGFR1c protein as discussed above. Stabilising mutations could be introduced to the FGFs such as those used to engineer long-acting FGF21 therapeutic biologic molecules (Table 4.3), however this would result in non-native protein and is a less favoured option. Finally, only freshly prepared FGF19 and FGF21 proteins should be used to ensure minimal degradation and maximal activity are maintained during experiments.

4.4.4 Recombinant βKlotho protein production – an experimental bottleneck?

In this work, the primary limitation was the lack of sufficient quantities of βKlotho protein expression for purification. This impeded experimental analyses of ternary signalling complexes *in vitro* including the much sought βKlotho and binary and ternary complex structural information. Although small quantities of extracellular and full length βKlotho could eventually be expressed (Figures 4.19 and 4.20), further optimisation of the detergent extraction and purification methods and large scale-ups will be needed before sufficient quantities of this protein can be robustly extracted from mammalian cells. In previous work, human βKlotho (1-992) was expressed and purified from HEK 293F cells (Yie et al., 2009). Mouse orthologues of αKlotho (35-982) and βKlotho (53-995) were also expressed and purified from HEK 293F cells in three studies performed by the same group (Goetz et al., 2012a, Goetz et al., 2012b, Goetz et al., 2010). A chimeric mouse αKlotho ectodomain with a C-terminal fusion to human IgG1 Fc (Tohyama et al., 2004) and the corresponding human αKlotho orthologue (Yamazaki et al., 2010) were also previously expressed and purified from the conditioned medium of CHO cells. In this study, it was not possible to detect similar secretion of βKlotho into the conditioned medium from CHO cells or HEK 293F cells, despite small amounts of intracellular expression being observed (Figure 4.19c and 4.19f). The constructs used in this work have slightly different N- and C-terminal boundaries (by a few amino acids) compared to some of the constructs used in the previous studies (Yie et al., 2009). Optimisation of the construct boundaries, secretion signal peptides and expression conditions in the mammalian systems used in this work should eventually lead to the successful secretion

of βKlotho ectodomain for purification. Much of the eagerly awaited molecular information required for a complete understanding of these complexes and their pathophysiological roles will continue to rely on the ability to produce sufficient quantities (and of a high purity) of the Klotho proteins.

4.4.5 Preliminary characterisation of binary endocrine FGF-FGFR complexes

With the protein components that were possible to produce recombinantly in this work, preliminary protein-protein binding experiments were performed using FGF19, FGF21 and the ectodomain of FGFR1c. It was initially attempted to analyse these interactions between FGF19-FGFR1c and FGF21-FGFR1c using ITC to probe the energetics, kinetics and stoichiometries of the binary ligand-receptor interactions. However, it was found to be challenging to use this technique to characterise FGF-FGFR interactions, partly due to the very small released heats of interaction (Figure 4.22). This led to attempts to reproduce the binary interaction data generated previously for FGF19-FGFR1c and FGF21-FGFR1c (Goetz et al., 2012b), albeit using a longer length and more physiologically relevant FGFR1c construct (D1-D3) than the published version (D2-D3, lacking the D1 domain). Preliminary SPR experiments yielded K_D values two orders of magnitude lower than those published in the above study (Table 4.6). These discrepancies between the two datasets could reflect several possibilities. Firstly, in this work, the FGFR1c construct used comprised the additional D1 domain whose role in binding to endocrine FGFs and/or Klotho-co-receptors has not been fully elucidated, as most biochemical and biophysical analyses have utilised the shorter length FGFR1c D2-D3 construct (discussed above). Additionally, the FGFR1c protein produced in this work was glycosylated due to production from mammalian cells (Figure 4.21), whereas the D2-D3 protein in the published study was refolded from bacterial inclusion bodies and so was not glycosylated. Since stronger binding between endocrine FGFs and FGFR1c was observed in this work using the glycosylated D1-D3 construct compared to the published study using the non-glycosylated D2-D3 construct (Table 4.6), this works tentatively suggests that the D1 domain and/or specific glycosyl moieties might play a role in the endocrine FGF-FGFR interaction, although previous studies have proposed roles for the D2 and D3 domains in the FGF-FGFR interaction (Goetz et al., 2012a, Gupte et al., 2011, Wu et al., 2011). Additionally, no studies have methodically investigated the importance of glycosylation in the binary and/or ternary complexes. Further experiments, including using a control paracrine FGF such as FGF1, and both glycosylated and non-glycosylated proteins, are needed to verify these data. Secondly, the likely heterogeneity of the refolded FGFs and analyte protein sticking to the SPR microfluidics added complexity to the analysis of these data, as evidenced by the shape of the steady-state binding curves (Figure 4.23). More robust preparations of protein and optimisation of the

SPR buffer conditions are required to improve the SPR data quality to extract reliable information about binding affinity and kinetics. In this work, it was not possible to fit the data to any kinetic model, and in fact little kinetic data has been reported for any endocrine FGF-FGFR interactions, with one study attempting to perform kinetic experiments with these complexes in cells (Yang et al., 2012). This could reflect the inherent complexity of the interactions, which might involve transient or constitutive homodimerisation of FGF ligands and/or FGFRs in solution in addition to the FGF-FGFR heterodimerisation that these experiments attempt to report on. Such biological complexity cannot be modelled by the 'simple' 1:1 or 2:1 kinetic models inbuilt into the SPR analytical software. In any case, these preliminary binding experiments provide a useful starting point to further probe these interactions. Further optimisations of these experiments will provide more accurate information about the true affinities and kinetics involved in binding of the endocrine FGFs to FGFR1c, and can eventually be extended to similar analyses of the ternary complexes involving Klotho proteins. This information will then eventually inform the clinical development of FGF21 and its biologic drug candidates for anti-diabetes and anti-obesity treatments, and of anti-FGF19 therapeutics in cancer.

4.5 Summary

In summary, this work laid the foundations for future experiments to study the protein-protein interactions within endocrine FGFR signalling complexes from a structural, biophysical and biochemical perspective. Homology models of the βKlotho KL domains were generated to aid in experimental planning and these will also aid structure-based drug design in the future in the absence of any Klotho atomic structures. Progress was made in recombinant protein production and quality control, with the indication that mammalian expression systems might perform better when attempting to produce large quantities of extracellular proteins for biophysical studies. Preliminary characterisations of protein-protein interactions involved in βKlotho endocrine signalling complexes were performed similarly to those performed by others with a slightly different construct, and possible explanations for deviations from the published study in the early results were suggested based on the slightly different experimental conditions used in this work. Methods and experiments tested and found to be unproductive will guide future studies based on progress established in this work. Overall, the priority for the future is production of βKlotho protein in sufficient quantities to allow determination of the much-awaited molecular information of ternary endocrine FGFR signalling complexes. This is needed to inform the rational development of therapeutics for targeting deregulated endocrine FGFR signalling involved in metabolic pathologies and cancer.

Chapter 5: Conclusions and Perspectives

In this thesis, two aspects of the diverse modes of FGFR signalling involved in cancer and several other adverse pathologies were investigated. In Chapter 3, a comprehensive panel of 25 missense substitutions identified in the FGFR KD was assembled and the variant KD proteins produced. The impacts of these cancer-associated variants on activation, catalysis, substrate binding affinity and stability of the FGFR KD were assessed using biochemical and biophysical *in vitro* assays. The most interesting mutations identified from *in vitro* assays were further investigated in model stable cell lines in the context of the full length receptor, and their effects on downstream signalling and transformation were assessed. In Chapter 4, several protein components of endocrine FGFR signalling complexes involving endocrine ligands FGF19 and FGF21, FGFR1c and the FGFR co-receptor β-Klotho were produced and preliminarily characterised. The early stages of a project aiming to investigate the protein-protein interactions in the ternary endocrine signalling complexes from structural, biophysical and biochemical perspectives was therefore established. In this Chapter, the key findings from this thesis will be summarised and discussed, including a synoptic discussion linking common concepts between the two subject areas and cancer. The importance of this study will be linked to the current status of cancer-related FGFR and RTK therapies undergoing clinical trials, stressing the importance of a progression towards a patient-personalised approach to targeted therapeutics in combating cancer.

5.1 Why study cancer-associated point mutations?

In Chapter 3, the importance of both genetic hotspot and non-hotspot mutations as oncogenic drivers in FGFRs was highlighted, with a focus on missense substitutions in cancer, although oncogenic fusion proteins and FGFR gene amplifications also play important roles in oncogenesis (Turner and Grose, 2010, Gallo et al., 2015). A similar phenomenon could apply to other kinases (and even other oncogenes and tumour suppressor genes) in cancer. In this work, known and new potential driver mutations in the FGFR KD were identified using a combination of *in vitro* and cell-based assays while several mutations were classified as probable passenger mutations, and some were hypothesised to be oncogenic in the optimal cellular environment. This work therefore contributes to the growing understanding of the oncogenic mechanisms and functions of oncogenic driver mutations in FGFRs, adding to knowledge for those previously identified as drivers in FGFR3 (section 3.1). Defining specific FGFR aberrations that are most likely to benefit from specific targeted therapies in the clinic will hopefully help to boost the success of candidate drugs currently in clinical trials. Studies using a

combination of complementary assays similar to those in this work could be employed for other kinases, to better link the currently biased genetic data in mutation databases to function. Particularly, this will be useful for kinases that are also drug targets for various diseases. Such interdisciplinary studies will help to validate new oncogenic driver mutations which can then be tested *in vivo* to confirm their oncogenicity and assess the oncogene dependency of tumours on individual aberrations. In addition, experimental datasets such as that generated in this work using large panels of mutations can help to improve the ability of computational algorithms to predict the pathogenicity of new gain-of-function mutations that are constantly being identified from the ongoing international cancer genome sequencing projects. This will ultimately accelerate the process by which new driver mutations can be pre-clinically experimentally validated as therapeutic targets to inform specific targeting in clinical trials.

5.2 Why study Klotho-dependent endocrine signalling complexes?

In Chapter 4, less-well researched endocrine FGFR signalling complexes involving the βKlotho FGFR co-receptor were preliminarily characterised from a biochemical and biophysical perspective. This study was performed due to endocrine signalling being deregulated in various metabolic disorders and cancer. In the case of metabolic diseases, the focus has so far been on developing FGF21 and related analogues as agonist treatments for a range of metabolic diseases including diabetes, obesity and ageing (Table 4.3). In cancer, *FGF19* amplification has been identified as the driver mutation for a subset of liver cancers for which it also confers sensitivity to FGFR-specific inhibition (Sawey et al., 2011, Guagnano et al., 2012), making this branch of endocrine signalling a relatively new therapeutic target.

Additionally, endocrine signalling is also found deregulated in response to targeted inhibition of oncogenic FGFR signalling in the clinic, due to inhibition of FGF23-αKlotho signalling (Brown et al., 2005, Nogova et al., 2017). Currently, anti-FGFR targeting using pan-FGFR inhibitors results in side-effects due to these on-target toxicities. These FGF23 signalling-related on-target toxicities arise because FGFR signalling regulates a plethora of cellular functions, so potent inhibition of oncogenic FGFR signalling also has knock-on adverse effects on normal 'beneficial' FGFR signalling (such as through FGF23) *in vivo*. Targeting these on-target adverse effects of potent pan-FGFR inhibitors therefore requires a molecular understanding of the mechanisms giving rise to these toxicities, which occur due to the co-occurrence of paracrine and endocrine signalling *in vivo*. Molecular studies of endocrine Klotho-dependent signalling complexes are therefore urgently needed to complement the existing wealth of information about

paracrine signalling, so that appropriate data-driven dosing regimens can be implemented in current clinical trials with potent pan-FGFR inhibitors. Molecular information will also aid the rational design of agonists to target and activate specific 'beneficial' signalling complexes alongside inhibition of canonical 'oncogenic' FGFR signalling, to help to alleviate the on-target toxicities caused by clinical pan-FGFR inhibition. One example of an agonist (effective against the FGFR1c-βKlotho complex, Table 4.3) has recently entered pre-clinical development (Sakamoto et al., 2016). However, in general, the Klotho-dependent sub-axes of FGFR signalling need to be better studied from a molecular perspective to rationally address several of these clinical problems.

5.3 Challenges and opportunities of targeting aberrant FGFR signalling

There are several challenges associated with targeting aberrant FGFR signalling in the clinic. As previously mentioned, the field of specific FGFR-targeting in the clinic has so far lagged behind other RTKs despite pre-clinical evidence implicating several aspects of FGFR signalling in oncogenesis. The importance of patient selection through identifying subsets of patients with specific FGFR aberrations for a personalised approach to treatment in the clinic has already been stressed in this work. In addition, the success of anti-FGFR targeting in the clinic has so far been variable despite pre-clinical models validating FGFRs as promising therapeutic targets. This is probably due to a combination of variable addiction of tumours to aberrant FGFR signalling and tumour heterogeneity (including heterogeneity of FGFR aberrations within a cancer) which might limit the efficacy of single-agent targeted therapies. Toxic effects of targeted therapies in the clinic currently result in withdrawal of treatments (including early termination of clinical trials) and/or dose reductions which limit the potential to eliminate a tumour. Finally, it is widely known that cancers evolve and respond to therapeutic inhibition by acquiring mechanisms of resistance to targeted therapies. These challenges associated with clinical FGFR targeting are discussed further below, emphasising the importance of this work, and highlighting aspects where further efforts should focus to improve the outcomes of personalised anti-FGFR therapy. These concepts are summarised in Figure 5.1.

5.3.1 Importance of patient selection based on specific FGFR aberrations

The first challenge of patient selection has already been discussed in this work. It is clear from both pre-clinical and clinical evidence that the importance of considering specific aberrations in FGFRs (and other oncogenes) is crucial for distinguishing between oncogenic driver and passenger mutations (Greenman et al., 2007, Guagnano et al.,

2012, Patani et al., 2016). In this work, it was shown that of the hundreds of mutations found in FGFRs, only a subset are activating and an even smaller subset are strongly transforming. An additional complication arises because distinct aberrations are only observed infrequently in the clinic, making it difficult to define stratified populations of patients that will benefit most from specific targeted therapies. Studies that functionally characterise large panels of cancer-associated mutations as in this work to distinguish between oncogenic driver and passenger mutations will help towards defining patient subsets that are most likely to present opportunities to respond to specific FGFR inhibitors in the clinic. This strategy must be undertaken alongside patient screening prior to treatment to identify the specific genetic aberrations within each tumour. This will help to provide treatment options with the best monotherapy or combinations of anti-FGFR therapies to boost the translation of the pre-clinical efficacy of candidate drugs to the clinic. Furthermore, an understanding of the mechanisms by which different mutations affect sensitivity to therapeutics can help to define different treatment options for patients based on their aberrations. For example, more specific inhibition of FGFR2 or FGFR3 using therapeutic antibodies which target the extracellular domains might be useful for tumours in which mutations in the KD alter conformation of the intracellular region and so reduce sensitivity to small molecule inhibitors targeting the KD, without affecting the ectodomains. These and more molecular selection approaches are recently beginning to be applied in the clinic (Andre et al., 2011).

5.3.2 Validating oncogene addiction status of a tumour to an FGFR aberration

Often pre-clinical identification of oncogenic driver mutations is insufficient to translate to clinical responses to targeted inhibition. An additional complication comes from growing evidence that only subsets of cancers harbouring FGFR aberrations are oncogene-addicted to the altered FGFR signalling, and it is currently difficult to predict which subsets these might be without *in vivo* data. Tumours that are not dependent on the aberrant FGFR signalling for proliferation and survival are less likely to respond to anti-FGFR treatments. Therefore, alongside identifying driver FGFR aberrations as in this work, the next important step should include validation of the oncogene addiction status of a tumour to its aberration(s). For example, despite *FGFR1* amplification being the most frequent aberration in FGFRs across tumours (Carter et al., 2015, Helsten et al., 2016) and demonstrating sensitivity to targeted therapies that is associated with amplification status in pre-clinical models (Shiang et al., 2010, Weiss et al., 2010, Guagnano et al., 2012), *in vivo* evidence suggests that FGFR1 is not a dominant oncogene (Babina and Turner, 2017). It has been suggested that higher *FGFR* copy-number amplification (and therefore overexpression), the downstream signalling pathways activated and co-aberrant genes present in the same tumour all affect

oncogene dependency, defining small subsets of patients likely to benefit from *FGFR1* targeting, such as in breast cancer (Bedussi et al., 2014).

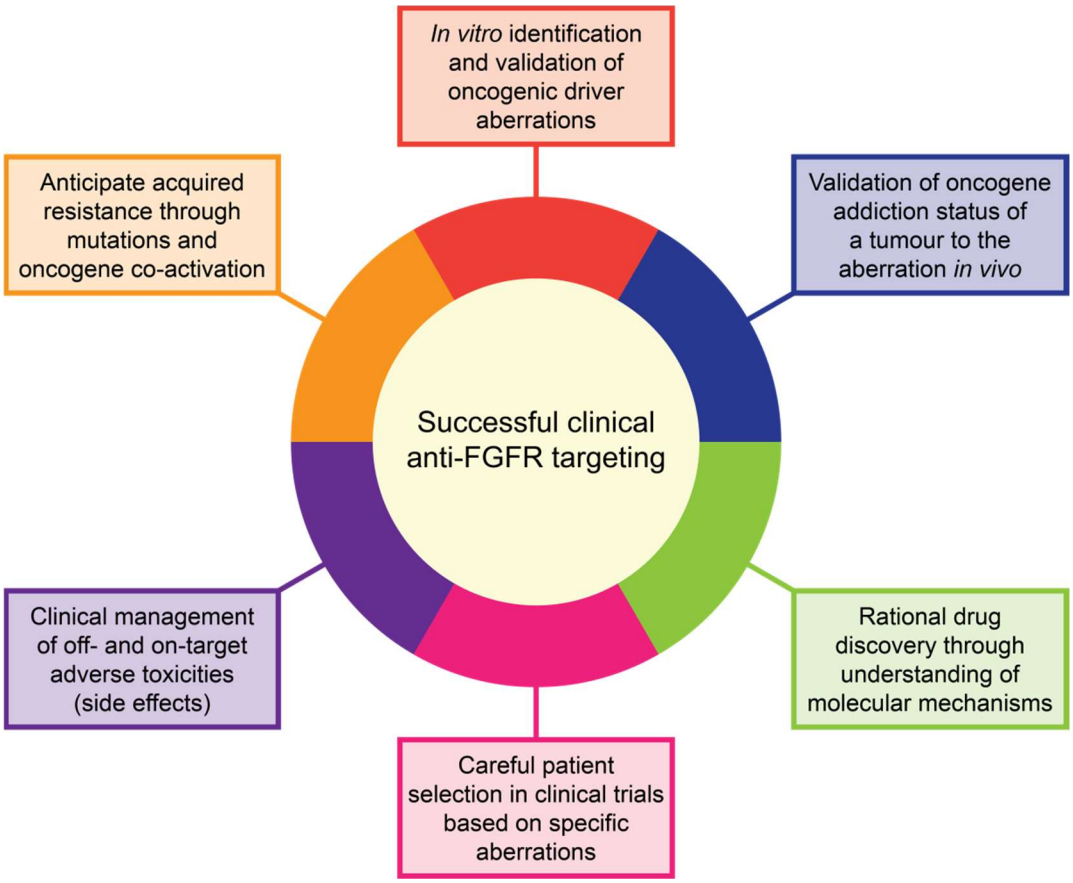


Figure 5.1 The challenges of clinical FGFR-targeting. A summary of the current challenges that need to be addressed for a future personalised approach to successful anti-FGFR targeting in the clinic. Addressing these areas should lead to therapies eventually being approved on the basis of anti-FGFR targeting. This thesis makes a novel contribution to some of the pre-clinical aspects of these challenges.

Conversely, *FGF19* amplification has been demonstrated to be an inhibitor-sensitive oncogenic driver aberration in pre-clinical hepatocellular carcinoma models, but not in other cancers (Sawey et al., 2011, Guagnano et al., 2012, Futami et al., 2017). This is dependent on the co-expression of βKlotho, and together these observations correlate with *in vivo* evidence showing that FGF19 overexpression is tumourigenic in mice (Ge et al., 2012). This provides a stronger case for the context dependency and oncogenicity of *FGF19* amplification in cancer. Although no clinical evidence is yet available, this early pre-clinical evidence suggests that tumours overexpressing FGF19 in the presence of βKlotho and FGFR4 are likely to define a subset of hepatocellular carcinomas likely to

benefit from specific targeting of the FGF19-FGFR4-βKlotho signalling complex. In contrast, complexes with FGFR1c-βKlotho are thought to regulate the metabolic roles of FGF19. Therefore, specific targeting of FGFR4 might produce less adverse side-effects than the current strategy of pan-FGFR targeting, by inhibiting only oncogenic effects of FGF19 signalling without affecting its 'beneficial' metabolic roles. Anti-FGFR4 inhibitors currently in pre-clinical development or in clinical trials are listed in Table 4.3. Two new clinical trials against hepatocellular carcinomas in patients specifically harbouring *FGF19* amplification (using the pan-FGFR inhibitor JNJ-42756493) or expressing FGFR4 and βKlotho (using a new FGFR4 inhibitor NVP-FGF401) are currently underway (Table 4.3). However, validation of the oncogene addiction status of FGF19-βKlotho signalling in the clinic is still required from these ongoing clinical trials to confirm the hypothesised roles of FGFR1c/FGFR4 in mediating the different FGF19 dependent signalling outcomes, which are currently controversial (section 4.1). Progress made towards the molecular characterisation of FGF19 endocrine signalling complexes as in this work with FGFR1c and in future studies with FGFR4 will aid the rational design of targeted agonists or antagonists against FGF19-βKlotho complexes in the different contexts of hepatocellular carcinomas and in metabolic diseases.

As a final example, despite the diversity of point mutations and oncogenic fusions found in FGFRs in cancer (particularly in FGFR3 and FGFR2), many of which are transforming and inhibitor-sensitive in pre-clinical models as shown in this work and by others (Tomlinson et al., 2007b, di Martino et al., 2009, Singh et al., 2012, Byron et al., 2013, Liao et al., 2013, Williams et al., 2013, Wu et al., 2013, Patani et al., 2016), not much is known about the specific effects of each mutation on oncogene addiction to FGFR signalling in tumours. Observations from several of the above studies suggest that tumours harbouring oncogenic fusions might confer greater inhibitor sensitivity than those harbouring individual point mutations, but individual point mutations also confer differing degrees of inhibitor sensitivity (as discussed in section 3.4). It is not known whether this is due to greater oncogene dependence of the tumour models on fusion-dependent signalling, or whether point mutations alter inhibitor efficacy compared to the fusion proteins which tend to have a 'WT' protein sequence (against which most drug candidates are validated), reducing inhibitor potency. Direct comparisons between the impacts of oncogenic fusions and point mutations on inhibitor sensitivity have rarely been made, and should be prioritised in the future. In this work, a large panel of FGFR3 KD mutations was ranked by their activation status and ability to transform cells, forming a pre-clinical basis for the identification of potential oncogenic driver point mutations. The addiction status of tumours to these point mutations and to all types of FGFR aberrations (including FGFR fusions) will need to be validated by data from the ongoing clinical trials.

5.3.3 The problem of on-target toxicities due to clinical pan-FGFR targeting

To date, there are still no approved therapies based on FGFR-targeting, but both multi-kinase and FGFR-specific (pan-FGFR) inhibitors are undergoing clinical trials on this basis. The results from these clinical studies revealed that inhibitors with different target specificities ultimately displayed varied toxicity profiles. Multi-kinase inhibitors that co-target other closely related RTKs such as VEGFRs, PDGFRs and KIT with nanomolar efficacy tended to exhibit off-target adverse effects due to co-inhibition of these other RTKs (Dieci et al., 2013, Shah et al., 2013b, Dienstmann et al., 2014, Soria et al., 2014). This led to rapid development of specific anti-FGFR targeted agents, of which many small molecule inhibitors target several FGFRs (most commonly FGFR1-3 which have the most similar KDs) (Table 1.4). However, due to the pleiotropism of FGFR signalling in all types of tissues and organs, potent FGFR-specific therapies were found to exhibit adverse on-target toxicities *in vivo*. Predominantly, these include hyperphosphatemia-mediated tissue calcification and bone mineralisation due to inhibition of FGF23-αKlotho endocrine signalling which regulates vitamin D and phosphate homeostasis (Brown et al., 2005, Nogova et al., 2017). Nevertheless, it is believed that these toxicities are manageable (Herbert et al., 2014) and that the more favourable toxicity profiles of pan-FGFR inhibitors suggest they are still overall better clinical candidates than multi-kinase inhibitors for FGFR targeting (Dieci et al., 2013, Dienstmann et al., 2014). However, studying oncogenic mutations will ultimately be required to define which inhibitors should be used for the most efficacious targeting of different subsets of FGFR aberrations. In order to limit knock-on inhibition of FGF23 signalling inhibitors, doses of potent FGFR inhibitors and inclusion of co-treatments for expected toxicities in treatment regimens should be carefully controlled. Alternatively, FGF-ligand traps such as the clinical candidate FP-1039 can be used, since early clinical data suggest that this molecule does not cause endocrine FGF-related toxicities (Herbert et al., 2014). Additionally, using FGFR2- or FGFR3-specific therapeutic antibodies (that have just entered early Phase I clinical trials, Table 1.4) for tumours harbouring aberrations of these genes might alleviate toxicities that potentially occur through FGF23-αKlotho-FGFR1 complexes (although this is yet to be validated clinically). Characterisation of FGF23-αKlotho signalling complexes to determine the FGFR isoform specificity of its phosphate homeostatic effects, and to aid in the design of new agonists targeting these complexes, will greatly help in management of the FGF23-dependent clinical toxicities, and will allow higher doses of FGFR inhibitors to be used in cancer targeting. Together, this will hopefully help to eliminate FGFR-aberrant tumours.

5.3.4 Tumour mechanisms of acquired resistance to targeted therapies

Currently, dosing limitations or even treatment withdrawal of potent FGFR inhibitors due to on-target toxicities contribute to the next challenge of clinical FGFR inhibition: the development of acquired resistance mutations by tumours. The observation of acquired gatekeeper mutations in FGFRs and other RTKs in response to targeted ATP-competitive inhibitors has already been discussed in detail in sections 1.1.2 and 3.1.6 as a major limitation of optimal efficacy for this class of FGFR inhibitors. The acquisition of resistance point mutations by tumours should continue to be anticipated as clinical trials progress, however it should be noted that gatekeeper FGFR mutations (and other mutations associated with inhibitor resistance) confer resistance to different degrees against individual inhibitors (Bunney et al., 2015, Byron et al., 2013, Patani et al., 2016, Chell et al., 2013). This work also suggests that FGFR3-N540K, the gatekeeper FGFR3-V555M and several other activating FGFR3 mutations likely impart different degrees of resistance to ATP-competitive inhibitors due to their increased affinities for ATP (Figure 3.14d). Those inhibitors that retain efficacy (based on pre-clinical testing in the above studies) could still be tested clinically in the event of resistance mutations in a tumour. Additionally, the recent development of second-generation irreversible inhibitors that bind covalently in the ATP-binding pocket through a P-loop cysteine residue conserved in FGFRs (Tan et al., 2014, Huang et al., 2015) are capable of maintaining their efficacy against gatekeeper and other resistance mutations in the ATP-binding pocket. Further clinical development of these molecules might help to provide second-line treatments against tumours refractory to currently used FGFR targeted therapies due to acquired resistance. *In vitro* assays as developed in this work can be used to identify mutations likely to confer resistance to ATP-competitive inhibitors and to screen new second-generation inhibitors against large panels of FGFR variants to validate their efficacies prior to clinical development of lead compounds.

Another recently observed mechanism of resistance in pre-clinical hepatocellular carcinoma models was to the multi-kinase inhibitor Sorafenib through activation of FGF19-FGFR4-βKlotho signalling, although these cells could be effectively targeted using a novel pan-FGFR inhibitor ASP5878 (Futami et al., 2017, Gao et al., 2017). Similar ‘escape’ mechanisms of resistance through targeting of another RTK, such as EGFR, have been shown to lead to resistance to the targeted therapy via activation of FGFR signalling (Crystal et al., 2014). Conversely, the opposite effect, whereby resistance to FGFR inhibitors occurs through activation of signalling via another RTK such as EGFR, has also been observed (Herrera-Abreu et al., 2013). Both mechanisms also occur with several other RTKs, already detailed in section 3.1.6, and this defines a different subset of clinical resistance mechanisms involving FGFRs and co-activation of

other oncogenes. These clear indications of crosstalk between RTKs that are co-activated in cancer and their observed strategies to co-operatively develop and maintain oncogenic signalling networks in the event of monotherapy against a single RTK highlight the emerging importance of dual- or multi-targeted inhibition in such cancers. In this work, preliminary data suggested there was indeed crosstalk between different signalling pathways activated in model FGFR3 point mutated cell lines. In these contexts, monotherapies will ultimately be ineffective and might even facilitate the development of acquired resistance through oncogene switching. This will be particularly dependent on whether the tumour is addicted to signalling via the RTK being targeted. Similarly, other non-RTK oncogenes and downstream signalling proteins have also been observed to be co-activated with FGFRs in specific cancers (di Martino et al., 2016). Therefore, as discussed in this Chapter, understanding the oncogenic status of specific FGFR aberrations, those of other oncogenes, and the addiction status of the tumour to each of these and whether they activate common downstream signalling pathways will pave the way to administration of optimal multi-kinase inhibitors or combinations of monotherapies against resilient tumours in the clinic.

5.4 Conclusions

In summary, this work studied the impacts of 25 cancer-associated FGFR KD mutations on kinase activation, kinetics, catalysis and stability of the KD *in vitro*, as well as on transformation and downstream signalling in a cellular model. It was shown that all hotspot mutations are not necessarily activating, while non-hotspot mutations might in fact be highly activating (such as the newly identified R669G). The context-dependency of the oncogenic nature of activating mutations was also hypothesised based on these studies. In parallel, early progress was made towards the expression and purification of proteins and the analysis of βKlotho dependent endocrine FGFR signalling complexes which are involved in oncogenesis of specific subsets of cancers and in the manifestation of adverse responses to targeted clinical FGFR inhibition. Finally, the importance of identifying specific molecular determinants driving cancers, in order to inform targeted clinical therapeutics, was discussed. Together, this work summarises and contributes to the current status of progress made towards personalised approaches to FGFR targeting which should have real applications for patients in the clinic in the near future.

Bibliography

ABBOSH, P. H., MCCONKEY, D. J. & PLIMACK, E. R. 2015. Targeting Signaling Transduction Pathways in Bladder Cancer. *Curr Oncol Rep,* 17**,** 58.

ABRAMOVITZ, L., RUBINEK, T., LIGUMSKY, H., BOSE, S., BARSHACK, I., AVIVI, C., KAUFMAN, B. & WOLF, I. 2011. KL1 internal repeat mediates klotho tumor suppressor activities and inhibits bFGF and IGF-I signaling in pancreatic cancer. *Clin Cancer Res,* 17**,** 4254-66.

ADAMS, A. C., CHENG, C. C., COSKUN, T. & KHARITONENKOV, A. 2012. FGF21 requires betaklotho to act in vivo. *PLoS One,* 7**,** e49977.

ADAMS, J. A. 2001. Kinetic and catalytic mechanisms of protein kinases. *Chem Rev,* 101**,** 2271-90.

ADAR, R., MONSONEGO-ORNAN, E., DAVID, P. & YAYON, A. 2002. Differential activation of cysteine-substitution mutants of fibroblast growth factor receptor 3 is determined by cysteine localization. *J Bone Miner Res,* 17**,** 860-8.

ADHR CONSORTIUM, T. 2000. Autosomal dominant hypophosphataemic rickets is associated with mutations in FGF23. *Nat Genet,* 26**,** 345-8.

AGAZIE, Y. M., MOVILLA, N., ISCHENKO, I. & HAYMAN, M. J. 2003. The phosphotyrosine phosphatase SHP2 is a critical mediator of transformation induced by the oncogenic fibroblast growth factor receptor 3. *Oncogene,* 22**,** 6909-18.

AHMED, Z., GEORGE, R., LIN, C. C., SUEN, K. M., LEVITT, J. A., SUHLING, K. & LADBURY, J. E. 2010. Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cell Signal,* 22**,** 23-33.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

ANDRE, F., DELALOGE, S. & SORIA, J. C. 2011. Biology-driven phase II trials: what is the optimal model for molecular selection? *J Clin Oncol,* 29**,** 1236-8.

ARNOLD, K., BORDOLI, L., KOPP, J. & SCHWEDE, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics,* 22**,** 195-201.

ASADA, M., SHINOMIYA, M., SUZUKI, M., HONDA, E., SUGIMOTO, R., IKEKITA, M. & IMAMURA, T. 2009. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochim Biophys Acta,* 1790**,** 40-8.

AUBERTIN, J., TOURPIN, S., JANOT, F., AHOMADEGBE, J. C. & RADVANYI, F. 2007. Analysis of fibroblast growth factor receptor 3 G697C mutation in oral squamous cell carcinomas. *Int J Cancer,* 120**,** 2058-9; author reply 2060.

AZAM, M. & DALEY, G. Q. 2006. Anticipating clinical resistance to target-directed agents : the BCR-ABL paradigm. *Mol Diagn Ther,* 10**,** 67-76.

AZAM, M., SEELIGER, M. A., GRAY, N. S., KURIYAN, J. & DALEY, G. Q. 2008. Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat Struct Mol Biol,* 15**,** 1109-18.

BABINA, I. S. & TURNER, N. C. 2017. Advances and challenges in targeting FGFR signalling in cancer. *Nat Rev Cancer,* 17**,** 318-332.

BADMAN, M. K., PISSIOS, P., KENNEDY, A. R., KOUKOS, G., FLIER, J. S. & MARATOS-FLIER, E. 2007. Hepatic fibroblast growth factor 21 is regulated by PPARalpha and is a key mediator of hepatic lipid metabolism in ketotic states. *Cell Metab,* 5**,** 426-37.

BAE, J. H., BOGGON, T. J., TOME, F., MANDIYAN, V., LAX, I. & SCHLESSINGER, J. 2010. Asymmetric receptor contact is required for tyrosine autophosphorylation of fibroblast growth factor receptor in living cells. *Proc Natl Acad Sci U S A,* 107**,** 2866-71.

BAE, J. H., LEW, E. D., YUZAWA, S., TOME, F., LAX, I. & SCHLESSINGER, J. 2009. The selectivity of receptor tyrosine kinase signaling is controlled by a secondary SH2 domain binding site. *Cell,* 138**,** 514-24.

BAE, J. H. & SCHLESSINGER, J. 2010. Asymmetric tyrosine kinase arrangements in activation or autophosphorylation of receptor tyrosine kinases. *Mol Cells,* 29**,** 443-8.

BARRETT, T., SURESH, C. G., TOLLEY, S. P., DODSON, E. J. & HUGHES, M. A. 1995. The crystal structure of a cyanogenic beta-glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure,* 3**,** 951-60.

BEDUSSI, F., BOTTINI, A., MEMO, M., FOX, S. B., SIGALA, S. & GENERALI, D. 2014. Targeting fibroblast growth factor receptor in breast cancer: a promise or a pitfall? *Expert Opin Ther Targets,* 18**,** 665-78.

BEENKEN, A., ELISEENKOVA, A. V., IBRAHIMI, O. A., OLSEN, S. K. & MOHAMMADI, M. 2012. Plasticity in interactions of fibroblast growth factor 1 (FGF1) N terminus with FGF receptors underlies promiscuity of FGF1. *J Biol Chem,* 287**,** 3067-78.
BEENKEN, A. & MOHAMMADI, M. 2009. The FGF family: biology, pathophysiology and therapy. *Nat Rev Drug Discov,* 8**,** 235-53.
BELLOSTA, P., IWAHORI, A., PLOTNIKOV, A. N., ELISEENKOVA, A. V., BASILICO, C. & MOHAMMADI, M. 2001. Identification of receptor and heparin binding sites in fibroblast growth factor 4 by structure-based mutagenesis. *Mol Cell Biol,* 21**,** 5946-57.
BELLUS, G. A., SPECTOR, E. B., SPEISER, P. W., WEAVER, C. A., GARBER, A. T., BRYKE, C. R., ISRAEL, J., ROSENGREN, S. S., WEBSTER, M. K., DONOGHUE, D. J. & FRANCOMANO, C. A. 2000. Distinct missense mutations of the FGFR3 lys650 codon modulate receptor kinase activation and the severity of the skeletal dysplasia phenotype. *Am J Hum Genet,* 67**,** 1411-21.
BELOV, A. A. & MOHAMMADI, M. 2012. Grb2, a double-edged sword of receptor tyrosine kinase signaling. *Sci Signal,* 5**,** pe49.
BELOV, A. A. & MOHAMMADI, M. 2013. Molecular mechanisms of fibroblast growth factor signaling in physiology and pathology. *Cold Spring Harb Perspect Biol,* 5.
BENET-PAGES, A., ORLIK, P., STROM, T. M. & LORENZ-DEPIEREUX, B. 2005. An FGF23 missense mutation causes familial tumoral calcinosis with hyperphosphatemia. *Hum Mol Genet,* 14**,** 385-90.
BERNARD-PIERROT, I., BRAMS, A., DUNOIS-LARDE, C., CAILLAULT, A., DIEZ DE MEDINA, S. G., CAPPELLEN, D., GRAFF, G., THIERY, J. P., CHOPIN, D., RICOL, D. & RADVANYI, F. 2006. Oncogenic properties of the mutated forms of fibroblast growth factor receptor 3b. *Carcinogenesis,* 27**,** 740-7.
BERNETT, M. J., SOMASUNDARAM, T. & BLABER, M. 2004. An atomic resolution structure for human fibroblast growth factor 1. *Proteins,* 57**,** 626-34.
BILLEREY, C., CHOPIN, D., AUBRIOT-LORTON, M. H., RICOL, D., GIL DIEZ DE MEDINA, S., VAN RHIJN, B., BRALET, M. P., LEFRERE-BELDA, M. A., LAHAYE, J. B., ABBOU, C. C., BONAVENTURE, J., ZAFRANI, E. S., VAN DER KWAST, T., THIERY, J. P. & RADVANYI, F. 2001. Frequent FGFR3 mutations in papillary non-invasive bladder (pTa) tumors. *Am J Pathol,* 158**,** 1955-9.
BLOCH, L., SINESHCHEKOVA, O., REICHENBACH, D., REISS, K., SAFTIG, P., KURO-O, M. & KAETHER, C. 2009. Klotho is a substrate for alpha-, beta- and gamma-secretase. *FEBS Lett,* 583**,** 3221-4.
BLUNT, A. G., LAWSHE, A., CUNNINGHAM, M. L., SETO, M. L., ORNITZ, D. M. & MACARTHUR, C. A. 1997. Overlapping expression and redundant activation of mesenchymal fibroblast growth factor (FGF) receptors by alternatively spliced FGF-8 ligands. *J Biol Chem,* 272**,** 3733-8.
BOCHAROV, E. V., LESOVOY, D. M., GONCHARUK, S. A., GONCHARUK, M. V., HRISTOVA, K. & ARSENIEV, A. S. 2013. Structure of FGFR3 transmembrane domain dimer: implications for signaling and human pathologies. *Structure,* 21**,** 2087-93.
BOOKOUT, A. L., DE GROOT, M. H., OWEN, B. M., LEE, S., GAUTRON, L., LAWRENCE, H. L., DING, X., ELMQUIST, J. K., TAKAHASHI, J. S., MANGELSDORF, D. J. & KLIEWER, S. A. 2013. FGF21 regulates metabolism and circadian behavior by acting on the nervous system. *Nat Med,* 19**,** 1147-52.
BOTTCHER, R. T. & NIEHRS, C. 2005. Fibroblast growth factor signaling during early vertebrate development. *Endocr Rev,* 26**,** 63-77.
BROOKS, A. N., KILGOUR, E. & SMITH, P. D. 2012. Molecular pathways: fibroblast growth factor signaling: a new therapeutic opportunity in cancer. *Clin Cancer Res,* 18**,** 1855-62.
BROWN, A. P., COURTNEY, C. L., KING, L. M., GROOM, S. C. & GRAZIANO, M. J. 2005. Cartilage dysplasia and tissue mineralization in the rat following administration of a FGF receptor tyrosine kinase inhibitor. *Toxicol Pathol,* 33**,** 449-55.
BUNNEY, T. D., WAN, S., THIYAGARAJAN, N., SUTTO, L., WILLIAMS, S. V., ASHFORD, P., KOSS, H., KNOWLES, M. A., GERVASIO, F. L., COVENEY, P. V. & KATAN, M. 2015. The Effect of Mutations on Drug Sensitivity and Kinase Activity of Fibroblast Growth Factor Receptors: A Combined Experimental and Theoretical Study. *EBioMedicine,* 2**,** 194-204.
BURMEISTER, W. P., COTTAZ, S., DRIGUEZ, H., IORI, R., PALMIERI, S. & HENRISSAT, B. 1997. The crystal structures of Sinapis alba myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure,* 5**,** 663-75.

BURMEISTER, W. P., COTTAZ, S., ROLLIN, P., VASELLA, A. & HENRISSAT, B. 2000. High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J Biol Chem,* 275**,** 39385-93.
BYRON, S. A., CHEN, H., WORTMANN, A., LOCH, D., GARTSIDE, M. G., DEHKHODA, F., BLAIS, S. P., NEUBERT, T. A., MOHAMMADI, M. & POLLOCK, P. M. 2013. The N550K/H mutations in FGFR2 confer differential resistance to PD173074, dovitinib, and ponatinib ATP-competitive inhibitors. *Neoplasia,* 15**,** 975-88.
CANALES, A., LOZANO, R., LOPEZ-MENDEZ, B., ANGULO, J., OJEDA, R., NIETO, P. M., MARTIN-LOMAS, M., GIMENEZ-GALLEGO, G. & JIMENEZ-BARBERO, J. 2006. Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue. *FEBS J,* 273**,** 4716-27.
CANCER GENOME ATLAS RESEARCH, N., WEINSTEIN, J. N., COLLISSON, E. A., MILLS, G. B., SHAW, K. R., OZENBERGER, B. A., ELLROTT, K., SHMULEVICH, I., SANDER, C. & STUART, J. M. 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet,* 45**,** 1113-20.
CAPPELLEN, D., DE OLIVEIRA, C., RICOL, D., DE MEDINA, S., BOURDIN, J., SASTRE-GARAU, X., CHOPIN, D., THIERY, J. P. & RADVANYI, F. 1999. Frequent activating mutations of FGFR3 in human bladder and cervix carcinomas. *Nat Genet,* 23**,** 18-20.
CARPENTER, G. & JI, Q. 1999. Phospholipase C-gamma as a signal-transducing element. *Exp Cell Res,* 253**,** 15-24.
CARTER, E. P., FEARON, A. E. & GROSE, R. P. 2015. Careless talk costs lives: fibroblast growth factor receptor signalling and the consequences of pathway malfunction. *Trends Cell Biol,* 25**,** 221-33.
CASCONE, T., HERYNK, M. H., XU, L., DU, Z., KADARA, H., NILSSON, M. B., OBORN, C. J., PARK, Y. Y., EREZ, B., JACOBY, J. J., LEE, J. S., LIN, H. Y., CIARDIELLO, F., HERBST, R. S., LANGLEY, R. R. & HEYMACH, J. V. 2011. Upregulated stromal EGFR and vascular remodeling in mouse xenograft models of angiogenesis inhibitor-resistant human lung adenocarcinoma. *J Clin Invest,* 121**,** 1313-28.
CHA, S. K., ORTEGA, B., KUROSU, H., ROSENBLATT, K. P., KURO, O. M. & HUANG, C. L. 2008. Removal of sialic acid involving Klotho causes cell-surface retention of TRPV5 channel via binding to galectin-1. *Proc Natl Acad Sci U S A,* 105**,** 9805-10.
CHANG, Q., HOEFS, S., VAN DER KEMP, A. W., TOPALA, C. N., BINDELS, R. J. & HOENDEROP, J. G. 2005. The beta-glucuronidase klotho hydrolyzes and activates the TRPV5 channel. *Science,* 310**,** 490-3.
CHEFETZ, I., HELLER, R., GALLI-TSINOPOULOU, A., RICHARD, G., WOLLNIK, B., INDELMAN, M., KOERBER, F., TOPAZ, O., BERGMAN, R., SPRECHER, E. & SCHOENAU, E. 2005. A novel homozygous missense mutation in FGF23 causes Familial Tumoral Calcinosis associated with disseminated visceral calcification. *Hum Genet,* 118**,** 261-6.
CHELL, V., BALMANNO, K., LITTLE, A. S., WILSON, M., ANDREWS, S., BLOCKLEY, L., HAMPSON, M., GAVINE, P. R. & COOK, S. J. 2013. Tumour cell responses to new fibroblast growth factor receptor tyrosine kinase inhibitors and identification of a gatekeeper mutation in FGFR3 as a mechanism of acquired resistance. *Oncogene,* 32**,** 3059-70.
CHELLAIAH, A. T., MCEWEN, D. G., WERNER, S., XU, J. & ORNITZ, D. M. 1994. Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in immunoglobulin-like domain III creates a receptor highly specific for acidic FGF/FGF-1. *J Biol Chem,* 269**,** 11620-7.
CHEN, C. D., PODVIN, S., GILLESPIE, E., LEEMAN, S. E. & ABRAHAM, C. R. 2007a. Insulin stimulates the cleavage and release of the extracellular domain of Klotho by ADAM10 and ADAM17. *Proc Natl Acad Sci U S A,* 104**,** 19796-801.
CHEN, F., SARABIPOUR, S. & HRISTOVA, K. 2013a. Multiple consequences of a single amino acid pathogenic RTK mutation: the A391E mutation in FGFR3. *PLoS One,* 8**,** e56521.
CHEN, H., HUANG, Z., DUTTA, K., BLAIS, S., NEUBERT, T. A., LI, X., COWBURN, D., TRAASETH, N. J. & MOHAMMADI, M. 2013b. Cracking the molecular origin of intrinsic tyrosine kinase activity through analysis of pathogenic gain-of-function mutations. *Cell Rep,* 4**,** 376-84.
CHEN, H., MA, J., LI, W., ELISEENKOVA, A. V., XU, C., NEUBERT, T. A., MILLER, W. T. & MOHAMMADI, M. 2007b. A molecular brake in the kinase hinge region regulates the activity of receptor tyrosine kinases. *Mol Cell,* 27**,** 717-30.
CHEN, H., MARSIGLIA, W. M., CHO, M. K., HUANG, Z., DENG, J., BLAIS, S. P., GAI, W., BHATTACHARYA, S., NEUBERT, T. A., TRAASETH, N. J. & MOHAMMADI, M. 2017.

Elucidation of a four-site allosteric network in fibroblast growth factor receptor tyrosine kinases. *Elife,* 6.
CHEN, H., XU, C. F., MA, J., ELISEENKOVA, A. V., LI, W., POLLOCK, P. M., PITTELOUD, N., MILLER, W. T., NEUBERT, T. A. & MOHAMMADI, M. 2008. A crystallographic snapshot of tyrosine trans-phosphorylation in action. *Proc Natl Acad Sci U S A,* 105**,** 19660-5.
CHEN, J., WILLIAMS, I. R., LEE, B. H., DUCLOS, N., HUNTLY, B. J., DONOGHUE, D. J. & GILLILAND, D. G. 2005. Constitutively activated FGFR3 mutants signal through PLCgamma-dependent and -independent pathways for hematopoietic transformation. *Blood,* 106**,** 328-37.
CHESI, M., BRENTS, L. A., ELY, S. A., BAIS, C., ROBBIANI, D. F., MESRI, E. A., KUEHL, W. M. & BERGSAGEL, P. L. 2001. Activated fibroblast growth factor receptor 3 is an oncogene that contributes to tumor progression in multiple myeloma. *Blood,* 97**,** 729-36.
CHESI, M., NARDINI, E., BRENTS, L. A., SCHROCK, E., RIED, T., KUEHL, W. M. & BERGSAGEL, P. L. 1997. Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. *Nat Genet,* 16**,** 260-4.
CHOI, M., MOSCHETTA, A., BOOKOUT, A. L., PENG, L., UMETANI, M., HOLMSTROM, S. R., SUINO-POWELL, K., XU, H. E., RICHARDSON, J. A., GERARD, R. D., MANGELSDORF, D. J. & KLIEWER, S. A. 2006. Identification of a hormonal basis for gallbladder filling. *Nat Med,* 12**,** 1253-5.
CHOI, Y. L., SODA, M., YAMASHITA, Y., UENO, T., TAKASHIMA, J., NAKAJIMA, T., YATABE, Y., TAKEUCHI, K., HAMADA, T., HARUTA, H., ISHIKAWA, Y., KIMURA, H., MITSUDOMI, T., TANIO, Y., MANO, H. & GROUP, A. L. K. L. C. S. 2010. EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. *N Engl J Med,* 363**,** 1734-9.
CIACCIO, M. F., WAGNER, J. P., CHUU, C. P., LAUFFENBURGER, D. A. & JONES, R. B. 2010. Systems analysis of EGF receptor signaling dynamics with microwestern arrays. *Nat Methods,* 7**,** 148-55.
COMPS-AGRAR, L., DUNSHEE, D. R., EATON, D. L. & SONODA, J. 2015. Unliganded fibroblast growth factor receptor 1 forms density-independent dimers. *J Biol Chem,* 290**,** 24166-77.
COOLS, J., DEANGELO, D. J., GOTLIB, J., STOVER, E. H., LEGARE, R. D., CORTES, J., KUTOK, J., CLARK, J., GALINSKY, I., GRIFFIN, J. D., CROSS, N. C., TEFFERI, A., MALONE, J., ALAM, R., SCHRIER, S. L., SCHMID, J., ROSE, M., VANDENBERGHE, P., VERHOEF, G., BOOGAERTS, M., WLODARSKA, I., KANTARJIAN, H., MARYNEN, P., COUTRE, S. E., STONE, R. & GILLILAND, D. G. 2003. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med,* 348**,** 1201-14.
CROSSLEY, P. H. & MARTIN, G. R. 1995. The mouse Fgf8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development,* 121**,** 439-51.
CRYSTAL, A. S., SHAW, A. T., SEQUIST, L. V., FRIBOULET, L., NIEDERST, M. J., LOCKERMAN, E. L., FRIAS, R. L., GAINOR, J. F., AMZALLAG, A., GRENINGER, P., LEE, D., KALSY, A., GOMEZ-CARABALLO, M., ELAMINE, L., HOWE, E., HUR, W., LIFSHITS, E., ROBINSON, H. E., KATAYAMA, R., FABER, A. C., AWAD, M. M., RAMASWAMY, S., MINO-KENUDSON, M., IAFRATE, A. J., BENES, C. H. & ENGELMAN, J. A. 2014. Patient-derived models of acquired resistance can identify effective drug combinations for cancer. *Science,* 346**,** 1480-6.
D'AVIS, P. Y., ROBERTSON, S. C., MEYER, A. N., BARDWELL, W. M., WEBSTER, M. K. & DONOGHUE, D. J. 1998. Constitutive activation of fibroblast growth factor receptor 3 by mutations responsible for the lethal skeletal dysplasia thanatophoric dysplasia type I. *Cell Growth Differ,* 9**,** 71-8.
DEL PICCOLO, N., PLACONE, J. & HRISTOVA, K. 2015. Effect of thanatophoric dysplasia type I mutations on FGFR3 dimerization. *Biophys J,* 108**,** 272-8.
DESNOYERS, L. R., PAI, R., FERRANDO, R. E., HOTZEL, K., LE, T., ROSS, J., CARANO, R., D'SOUZA, A., QING, J., MOHTASHEMI, I., ASHKENAZI, A. & FRENCH, D. M. 2008. Targeting FGF19 inhibits tumor growth in colon cancer xenograft and FGF19 transgenic hepatocellular carcinoma models. *Oncogene,* 27**,** 85-97.
DI MARTINO, E., L'HOTE, C. G., KENNEDY, W., TOMLINSON, D. C. & KNOWLES, M. A. 2009. Mutant fibroblast growth factor receptor 3 induces intracellular signaling and

cellular transformation in a cell type- and mutation-specific manner. *Oncogene,* 28**,** 4306-16.
DI MARTINO, E., TOMLINSON, D. C., WILLIAMS, S. V. & KNOWLES, M. A. 2016. A place for precision medicine in bladder cancer: targeting the FGFRs. *Future Oncol,* 12**,** 2243-63.
DIECI, M. V., ARNEDOS, M., ANDRE, F. & SORIA, J. C. 2013. Fibroblast growth factor receptor inhibitors as a cancer treatment: from a biologic rationale to medical perspectives. *Cancer Discov,* 3**,** 264-79.
DIENSTMANN, R., RODON, J., PRAT, A., PEREZ-GARCIA, J., ADAMO, B., FELIP, E., CORTES, J., IAFRATE, A. J., NUCIFORO, P. & TABERNERO, J. 2014. Genomic aberrations in the FGFR pathway: opportunities for targeted therapies in solid tumors. *Ann Oncol,* 25**,** 552-63.
DING, X., BONEY-MONTOYA, J., OWEN, B. M., BOOKOUT, A. L., COATE, K. C., MANGELSDORF, D. J. & KLIEWER, S. A. 2012. betaKlotho is required for fibroblast growth factor 21 effects on growth and metabolism. *Cell Metab,* 16**,** 387-93.
DRUKER, B. J., TAMURA, S., BUCHDUNGER, E., OHNO, S., SEGAL, G. M., FANNING, S., ZIMMERMANN, J. & LYDON, N. B. 1996. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med,* 2**,** 561-6.
ELZI, D. J., SONG, M., BLACKMAN, B., WEINTRAUB, S. T., LOPEZ-TERRADA, D., CHEN, Y., TOMLINSON, G. E. & SHIIO, Y. 2016. FGF19 functions as autocrine growth factor for hepatoblastoma. *Genes Cancer,* 7**,** 125-35.
ENDRES, N. F., DAS, R., SMITH, A. W., ARKHIPOV, A., KOVACS, E., HUANG, Y., PELTON, J. G., SHAN, Y., SHAW, D. E., WEMMER, D. E., GROVES, J. T. & KURIYAN, J. 2013. Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell,* 152**,** 543-56.
EZZAT, S., HUANG, P., DACKIW, A. & ASA, S. L. 2005. Dual inhibition of RET and FGFR4 restrains medullary thyroid cancer cell growth. *Clin Cancer Res,* 11**,** 1336-41.
EZZAT, S., ZHENG, L., YU, S. & ASA, S. L. 2001. A soluble dominant negative fibroblast growth factor receptor 4 isoform in human MCF-7 breast cancer cells. *Biochem Biophys Res Commun,* 287**,** 60-5.
FENG, S., DAKHOVA, O., CREIGHTON, C. J. & ITTMANN, M. 2013. Endocrine fibroblast growth factor FGF19 promotes prostate cancer progression. *Cancer Res,* 73**,** 2551-62.
FENG, S., WANG, J., ZHANG, Y., CREIGHTON, C. J. & ITTMANN, M. 2015. FGF23 promotes prostate cancer progression. *Oncotarget,* 6**,** 17291-301.
FINCH, P. W., RUBIN, J. S., MIKI, T., RON, D. & AARONSON, S. A. 1989. Human KGF is FGF-related with properties of a paracrine effector of epithelial cell growth. *Science,* 245**,** 752-5.
FOLTZ, I. N., HU, S., KING, C., WU, X., YANG, C., WANG, W., WEISZMANN, J., STEVENS, J., CHEN, J. S., NUANMANEE, N., GUPTE, J., KOMOROWSKI, R., SEKIROV, L., HAGER, T., ARORA, T., GE, H., BARIBAULT, H., WANG, F., SHENG, J., KAROW, M., WANG, M., LUO, Y., MCKEEHAN, W., WANG, Z., VENIANT, M. M. & LI, Y. 2012. Treating diabetes and obesity with an FGF21-mimetic antibody activating the betaKlotho/FGFR1c receptor complex. *Sci Transl Med,* 4**,** 162ra153.
FON TACER, K., BOOKOUT, A. L., DING, X., KUROSU, H., JOHN, G. B., WANG, L., GOETZ, R., MOHAMMADI, M., KURO-O, M., MANGELSDORF, D. J. & KLIEWER, S. A. 2010. Research resource: Comprehensive expression atlas of the fibroblast growth factor system in adult mouse. *Mol Endocrinol,* 24**,** 2050-64.
FORBES, S. A., BEARE, D., GUNASEKARAN, P., LEUNG, K., BINDAL, N., BOUTSELAKIS, H., DING, M., BAMFORD, S., COLE, C., WARD, S., KOK, C. Y., JIA, M., DE, T., TEAGUE, J. W., STRATTON, M. R., MCDERMOTT, U. & CAMPBELL, P. J. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res,* 43**,** D805-11.
FRENCH, D. M., LIN, B. C., WANG, M., ADAMS, C., SHEK, T., HOTZEL, K., BOLON, B., FERRANDO, R., BLACKMORE, C., SCHROEDER, K., RODRIGUEZ, L. A., HRISTOPOULOS, M., VENOOK, R., ASHKENAZI, A. & DESNOYERS, L. R. 2012. Targeting FGFR4 inhibits hepatocellular carcinoma in preclinical mouse models. *PLoS One,* 7**,** e36713.
FROST, M. J., FERRAO, P. T., HUGHES, T. P. & ASHMAN, L. K. 2002. Juxtamembrane mutant V560GKit is more sensitive to Imatinib (STI571) compared with wild-type c-kit whereas the kinase domain mutant D816VKit is resistant. *Mol Cancer Ther,* 1**,** 1115-24.
FU, L., JOHN, L. M., ADAMS, S. H., YU, X. X., TOMLINSON, E., RENZ, M., WILLIAMS, P. M., SORIANO, R., CORPUZ, R., MOFFAT, B., VANDLEN, R., SIMMONS, L., FOSTER, J., STEPHAN, J. P., TSAI, S. P. & STEWART, T. A. 2004. Fibroblast growth factor 19

increases metabolic rate and reverses dietary and leptin-deficient diabetes. *Endocrinology,* 145**,** 2594-603.
FURDUI, C. M., LEW, E. D., SCHLESSINGER, J. & ANDERSON, K. S. 2006. Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol Cell,* 21**,** 711-7.
FURITSU, T., TSUJIMURA, T., TONO, T., IKEDA, H., KITAYAMA, H., KOSHIMIZU, U., SUGAHARA, H., BUTTERFIELD, J. H., ASHMAN, L. K., KANAYAMA, Y. & ET AL. 1993. Identification of mutations in the coding sequence of the proto-oncogene c-kit in a human mast cell leukemia cell line causing ligand-independent activation of c-kit product. *J Clin Invest,* 92**,** 1736-44.
FUTAMI, T., OKADA, H., KIHARA, R., KAWASE, T., NAKAYAMA, A., SUZUKI, T., KAMEDA, M., SHINDOH, N., TERASAKA, T., HIRANO, M. & KUROMITSU, S. 2017. ASP5878, a Novel Inhibitor of FGFR1, 2, 3, and 4, Inhibits the Growth of FGF19-Expressing Hepatocellular Carcinoma. *Mol Cancer Ther,* 16**,** 68-75.
GALLO, L. H., NELSON, K. N., MEYER, A. N. & DONOGHUE, D. J. 2015. Functions of Fibroblast Growth Factor Receptors in cancer defined by novel translocations and mutations. *Cytokine Growth Factor Rev,* 26**,** 425-49.
GAO, L., WANG, X., TANG, Y., HUANG, S., HU, C. A. & TENG, Y. 2017. FGF19/FGFR4 signaling contributes to the resistance of hepatocellular carcinoma to sorafenib. *J Exp Clin Cancer Res,* 36**,** 8.
GARCIA, D. & SHAW, R. J. 2017. AMPK: Mechanisms of Cellular Energy Sensing and Restoration of Metabolic Balance. *Mol Cell,* 66**,** 789-800.
GARRINGER, H. J., MALEKPOUR, M., ESTEGHAMAT, F., MORTAZAVI, S. M., DAVIS, S. I., FARROW, E. G., YU, X., ARKING, D. E., DIETZ, H. C. & WHITE, K. E. 2008. Molecular genetic and biochemical analyses of FGF23 mutations in familial tumoral calcinosis. *Am J Physiol Endocrinol Metab,* 295**,** E929-37.
GARTSIDE, M. G., CHEN, H., IBRAHIMI, O. A., BYRON, S. A., CURTIS, A. V., WELLENS, C. L., BENGSTON, A., YUDT, L. M., ELISEENKOVA, A. V., MA, J., CURTIN, J. A., HYDER, P., HARPER, U. L., RIEDESEL, E., MANN, G. J., TRENT, J. M., BASTIAN, B. C., MELTZER, P. S., MOHAMMADI, M. & POLLOCK, P. M. 2009. Loss-of-function fibroblast growth factor receptor-2 mutations in melanoma. *Mol Cancer Res,* 7**,** 41-54.
GE, H., BARIBAULT, H., VONDERFECHT, S., LEMON, B., WEISZMANN, J., GARDNER, J., LEE, K. J., GUPTE, J., MOOKHERJEE, P., WANG, M., SHENG, J., WU, X. & LI, Y. 2012. Characterization of a FGF19 variant with altered receptor specificity revealed a central role for FGFR1c in the regulation of glucose metabolism. *PLoS One,* 7**,** e33603.
GEMEL, J., GORRY, M., EHRLICH, G. D. & MACARTHUR, C. A. 1996. Structure and sequence of human FGF8. *Genomics,* 35**,** 253-7.
GOETZ, R., BEENKEN, A., IBRAHIMI, O. A., KALININA, J., OLSEN, S. K., ELISEENKOVA, A. V., XU, C., NEUBERT, T. A., ZHANG, F., LINHARDT, R. J., YU, X., WHITE, K. E., INAGAKI, T., KLIEWER, S. A., YAMAMOTO, M., KUROSU, H., OGAWA, Y., KURO-O, M., LANSKE, B., RAZZAQUE, M. S. & MOHAMMADI, M. 2007. Molecular insights into the klotho-dependent, endocrine mode of action of fibroblast growth factor 19 subfamily members. *Mol Cell Biol,* 27**,** 3417-28.
GOETZ, R. & MOHAMMADI, M. 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nat Rev Mol Cell Biol,* 14**,** 166-80.
GOETZ, R., NAKADA, Y., HU, M. C., KUROSU, H., WANG, L., NAKATANI, T., SHI, M., ELISEENKOVA, A. V., RAZZAQUE, M. S., MOE, O. W., KURO-O, M. & MOHAMMADI, M. 2010. Isolated C-terminal tail of FGF23 alleviates hypophosphatemia by inhibiting FGF23-FGFR-Klotho complex formation. *Proc Natl Acad Sci U S A,* 107**,** 407-12.
GOETZ, R., OHNISHI, M., DING, X., KUROSU, H., WANG, L., AKIYOSHI, J., MA, J., GAI, W., SIDIS, Y., PITTELOUD, N., KURO, O. M., RAZZAQUE, M. S. & MOHAMMADI, M. 2012a. Klotho coreceptors inhibit signaling by paracrine fibroblast growth factor 8 subfamily ligands. *Mol Cell Biol,* 32**,** 1944-54.
GOETZ, R., OHNISHI, M., KIR, S., KUROSU, H., WANG, L., PASTOR, J., MA, J., GAI, W., KURO-O, M., RAZZAQUE, M. S. & MOHAMMADI, M. 2012b. Conversion of a paracrine fibroblast growth factor into an endocrine fibroblast growth factor. *J Biol Chem,* 287**,** 29134-46.
GOLDFARB, M., SCHOORLEMMER, J., WILLIAMS, A., DIWAKAR, S., WANG, Q., HUANG, X., GIZA, J., TCHETCHIK, D., KELLEY, K., VEGA, A., MATTHEWS, G., ROSSI, P., ORNITZ, D. M. & D'ANGELO, E. 2007. Fibroblast growth factor homologous factors control neuronal excitability through modulation of voltage-gated sodium channels. *Neuron,* 55**,** 449-63.

GORRE, M. E., MOHAMMED, M., ELLWOOD, K., HSU, N., PAQUETTE, R., RAO, P. N. & SAWYERS, C. L. 2001. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science,* 293**,** 876-80.
GOSPODAROWICZ, D. & CHENG, J. 1986. Heparin protects basic and acidic FGF from inactivation. *J Cell Physiol,* 128**,** 475-84.
GREENMAN, C., STEPHENS, P., SMITH, R., DALGLIESH, G. L., HUNTER, C., BIGNELL, G., DAVIES, H., TEAGUE, J., BUTLER, A., STEVENS, C., EDKINS, S., O'MEARA, S., VASTRIK, I., SCHMIDT, E. E., AVIS, T., BARTHORPE, S., BHAMRA, G., BUCK, G., CHOUDHURY, B., CLEMENTS, J., COLE, J., DICKS, E., FORBES, S., GRAY, K., HALLIDAY, K., HARRISON, R., HILLS, K., HINTON, J., JENKINSON, A., JONES, D., MENZIES, A., MIRONENKO, T., PERRY, J., RAINE, K., RICHARDSON, D., SHEPHERD, R., SMALL, A., TOFTS, C., VARIAN, J., WEBB, T., WEST, S., WIDAA, S., YATES, A., CAHILL, D. P., LOUIS, D. N., GOLDSTRAW, P., NICHOLSON, A. G., BRASSEUR, F., LOOIJENGA, L., WEBER, B. L., CHIEW, Y. E., DEFAZIO, A., GREAVES, M. F., GREEN, A. R., CAMPBELL, P., BIRNEY, E., EASTON, D. F., CHENEVIX-TRENCH, G., TAN, M. H., KHOO, S. K., TEH, B. T., YUEN, S. T., LEUNG, S. Y., WOOSTER, R., FUTREAL, P. A. & STRATTON, M. R. 2007. Patterns of somatic mutation in human cancer genomes. *Nature,* 446**,** 153-8.
GUAGNANO, V., KAUFFMANN, A., WOHRLE, S., STAMM, C., ITO, M., BARYS, L., PORNON, A., YAO, Y., LI, F., ZHANG, Y., CHEN, Z., WILSON, C. J., BORDAS, V., LE DOUGET, M., GAITHER, L. A., BORAWSKI, J., MONAHAN, J. E., VENKATESAN, K., BRUMMENDORF, T., THOMAS, D. M., GARCIA-ECHEVERRIA, C., HOFMANN, F., SELLERS, W. R. & GRAUS-PORTA, D. 2012. FGFR genetic alterations predict for sensitivity to NVP-BGJ398, a selective pan-FGFR inhibitor. *Cancer Discov,* 2**,** 1118-33.
GUPTE, J., YANG, L., WU, X., WEISZMANN, J., HECHT, R., LEMON, B., LINDBERG, R., WANG, Z. & LI, Y. 2011. The FGFR D3 domain determines receptor selectivity for fibroblast growth factor 21. *J Mol Biol,* 408**,** 491-502.
HAFNER, C., DI MARTINO, E., PITT, E., STEMPFL, T., TOMLINSON, D., HARTMANN, A., LANDTHALER, M., KNOWLES, M. & VOGT, T. 2010. FGFR3 mutation affects cell growth, apoptosis and attachment in keratinocytes. *Exp Cell Res,* 316**,** 2008-16.
HAGE, N., RENSHAW, J. G., WINKLER, G. S., GELLERT, P., STOLNIK, S. & FALCONE, F. H. 2015. Improved expression and purification of the Helicobacter pylori adhesin BabA through the incorporation of a hexa-lysine tag. *Protein Expr Purif,* 106**,** 25-30.
HAGEL, M., MIDUTURU, C., SHEETS, M., RUBIN, N., WENG, W., STRANSKY, N., BIFULCO, N., KIM, J. L., HODOUS, B., BROOIJMANS, N., SHUTES, A., WINTER, C., LENGAUER, C., KOHL, N. E. & GUZI, T. 2015. First Selective Small Molecule Inhibitor of FGFR4 for the Treatment of Hepatocellular Carcinomas with an Activated FGFR4 Signaling Pathway. *Cancer Discov,* 5**,** 424-37.
HANKER, A. B., GARRETT, J. T., ESTRADA, M. V., MOORE, P. D., ERICSSON, P. G., KOCH, J. P., LANGLEY, E., SINGH, S., KIM, P. S., FRAMPTON, G. M., SANFORD, E., OWENS, P., BECKER, J., GROSECLOSE, M. R., CASTELLINO, S., JOENSUU, H., HUOBER, J., BRASE, J. C., MAJJAJ, S., BROHEE, S., VENET, D., BROWN, D., BASELGA, J., PICCART, M., SOTIRIOU, C. & ARTEAGA, C. L. 2017. HER2-Overexpressing Breast Cancers Amplify FGFR Signaling upon Acquisition of Resistance to Dual Therapeutic Blockade of HER2. *Clin Cancer Res,* 23**,** 4323-4334.
HANKS, S. K. & HUNTER, T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J,* 9**,** 576-96.
HARADA, M., MURAKAMI, H., OKAWA, A., OKIMOTO, N., HIRAOKA, S., NAKAHARA, T., AKASAKA, R., SHIRAISHI, Y., FUTATSUGI, N., MIZUTANI-KOSEKI, Y., KUROIWA, A., SHIROUZU, M., YOKOYAMA, S., TAIJI, M., ISEKI, S., ORNITZ, D. M. & KOSEKI, H. 2009. FGF9 monomer-dimer equilibrium regulates extracellular matrix affinity and tissue diffusion. *Nat Genet,* 41**,** 289-98.
HARMER, N. J., PELLEGRINI, L., CHIRGADZE, D., FERNANDEZ-RECIO, J. & BLUNDELL, T. L. 2004. The crystal structure of fibroblast growth factor (FGF) 19 reveals novel features of the FGF family and offers a structural basis for its unusual receptor affinity. *Biochemistry,* 43**,** 629-40.
HART, K. C., ROBERTSON, S. C., KANEMITSU, M. Y., MEYER, A. N., TYNAN, J. A. & DONOGHUE, D. J. 2000. Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4. *Oncogene,* 19**,** 3309-20.
HASHIMOTO, K., TSUJIMURA, T., MORIYAMA, Y., YAMATODANI, A., KIMURA, M., TOHYA, K., MORIMOTO, M., KITAYAMA, H., KANAKURA, Y. & KITAMURA, Y. 1996. Transforming and differentiation-inducing potential of constitutively activated c-kit

mutant genes in the IC-2 murine interleukin-3-dependent mast cell line. *Am J Pathol,* 148**,** 189-200.
HAUGSTEN, E. M., MALECKI, J., BJORKLUND, S. M., OLSNES, S. & WESCHE, J. 2008. Ubiquitination of fibroblast growth factor receptor 1 is required for its intracellular sorting but not for its endocytosis. *Mol Biol Cell,* 19**,** 3390-403.
HAUGSTEN, E. M., SORENSEN, V., BRECH, A., OLSNES, S. & WESCHE, J. 2005. Different intracellular trafficking of FGF1 endocytosed by the four homologous FGF receptors. *J Cell Sci,* 118**,** 3869-81.
HAUGSTEN, E. M., ZAKRZEWSKA, M., BRECH, A., PUST, S., OLSNES, S., SANDVIG, K. & WESCHE, J. 2011. Clathrin- and dynamin-independent endocytosis of FGFR3--implications for signalling. *PLoS One,* 6**,** e21708.
HAYASHI, Y., OKINO, N., KAKUTA, Y., SHIKANAI, T., TANI, M., NARIMATSU, H. & ITO, M. 2007. Klotho-related protein is a novel cytosolic neutral beta-glycosylceramidase. *J Biol Chem,* 282**,** 30889-900.
HECHT, R., LI, Y. S., SUN, J., BELOUSKI, E., HALL, M., HAGER, T., YIE, J., WANG, W., WINTERS, D., SMITH, S., SPAHR, C., TAM, L. T., SHEN, Z., STANISLAUS, S., CHINOOKOSWONG, N., LAU, Y., SICKMIER, A., MICHAELS, M. L., BOONE, T., VENIANT, M. M. & XU, J. 2012. Rationale-Based Engineering of a Potent Long-Acting FGF21 Analog for the Treatment of Type 2 Diabetes. *PLoS One,* 7**,** e49345.
HEINRICH, M. C., CORLESS, C. L., DUENSING, A., MCGREEVEY, L., CHEN, C. J., JOSEPH, N., SINGER, S., GRIFFITH, D. J., HALEY, A., TOWN, A., DEMETRI, G. D., FLETCHER, C. D. & FLETCHER, J. A. 2003. PDGFRA activating mutations in gastrointestinal stromal tumors. *Science,* 299**,** 708-10.
HELSTEN, T., ELKIN, S., ARTHUR, E., TOMSON, B. N., CARTER, J. & KURZROCK, R. 2016. The FGFR Landscape in Cancer: Analysis of 4,853 Tumors by Next-Generation Sequencing. *Clin Cancer Res,* 22**,** 259-67.
HELSTEN, T., SCHWAEDERLE, M. & KURZROCK, R. 2015. Fibroblast growth factor receptor signaling in hereditary and neoplastic disease: biologic and clinical implications. *Cancer Metastasis Rev,* 34**,** 479-96.
HENRISSAT, B. & BAIROCH, A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J,* 293 (Pt 3)**,** 781-8.
HERBERT, C., LASSALLE, G., ALCOUFFE, C. & BONO, F. 2014. Approaches targeting the FGF-FGFR system: a review of the recent patent literature and associated advanced therapeutic agents. *Pharm Pat Anal,* 3**,** 585-612.
HERRERA-ABREU, M. T., PEARSON, A., CAMPBELL, J., SHNYDER, S. D., KNOWLES, M. A., ASHWORTH, A. & TURNER, N. C. 2013. Parallel RNA interference screens identify EGFR activation as an escape mechanism in FGFR3-mutant cancer. *Cancer Discov,* 3**,** 1058-71.
HOLDMAN, X. B., WELTE, T., RAJAPAKSHE, K., POND, A., COARFA, C., MO, Q., HUANG, S., HILSENBECK, S. G., EDWARDS, D. P., ZHANG, X. & ROSEN, J. M. 2015. Upregulation of EGFR signaling is correlated with tumor stroma remodeling and tumor recurrence in FGFR1-driven breast cancer. *Breast Cancer Res,* 17**,** 141.
HU, L. & CONG, L. 2015. Fibroblast growth factor 19 is correlated with an unfavorable prognosis and promotes progression by activating fibroblast growth factor receptor 4 in advanced-stage serous ovarian cancer. *Oncol Rep,* 34**,** 2683-91.
HU, M. C., KURO-O, M. & MOE, O. W. 2012. Secreted klotho and chronic kidney disease. *Adv Exp Med Biol,* 728**,** 126-57.
HU, M. C., KURO-O, M. & MOE, O. W. 2013. Klotho and chronic kidney disease. *Contrib Nephrol,* 180**,** 47-63.
HU, M. C. & MOE, O. W. 2012. Klotho as a potential biomarker and therapy for acute kidney injury. *Nat Rev Nephrol,* 8**,** 423-9.
HU, M. C., SHI, M., ZHANG, J., PASTOR, J., NAKATANI, T., LANSKE, B., RAZZAQUE, M. S., ROSENBLATT, K. P., BAUM, M. G., KURO-O, M. & MOE, O. W. 2010. Klotho: a novel phosphaturic substance acting as an autocrine enzyme in the renal proximal tubule. *FASEB J,* 24**,** 3438-50.
HUANG, J., ISHINO, T., CHEN, G., ROLZIN, P., OSOTHPRAROP, T. F., RETTING, K., LI, L., JIN, P., MATIN, M. J., HUYGHE, B., TALUKDAR, S., BRADSHAW, C. W., PALANKI, M., VIOLAND, B. N., WOODNUTT, G., LAPPE, R. W., OGILVIE, K. & LEVIN, N. 2013a. Development of a novel long-acting antidiabetic FGF21 mimetic by targeted conjugation to a scaffold antibody. *J Pharmacol Exp Ther,* 346**,** 270-80.

HUANG, Z., CHEN, H., BLAIS, S., NEUBERT, T. A., LI, X. & MOHAMMADI, M. 2013b. Structural mimicry of a-loop tyrosine phosphorylation by a pathogenic FGF receptor 3 mutation. *Structure,* 21**,** 1889-96.

HUANG, Z., TAN, L., WANG, H., LIU, Y., BLAIS, S., DENG, J., NEUBERT, T. A., GRAY, N. S., LI, X. & MOHAMMADI, M. 2015. DFG-out mode of inhibition by an irreversible type-1 inhibitor capable of overcoming gate-keeper mutations in FGF receptors. *ACS Chem Biol,* 10**,** 299-309.

HUANG, Z., WANG, H., LU, M., SUN, C., WU, X., TAN, Y., YE, C., ZHU, G., WANG, X., CAI, L. & LI, X. 2011. A better anti-diabetic recombinant human fibroblast growth factor 21 (rhFGF21) modified with polyethylene glycol. *PLoS One,* 6**,** e20669.

HUBBARD, S. R. 1999. Structural analysis of receptor tyrosine kinases. *Prog Biophys Mol Biol,* 71**,** 343-58.

IBRAHIMI, O. A., ZHANG, F., ELISEENKOVA, A. V., LINHARDT, R. J. & MOHAMMADI, M. 2004. Proline to arginine mutations in FGF receptors 1 and 3 result in Pfeiffer and Muenke craniosynostosis syndromes through enhancement of FGF binding affinity. *Hum Mol Genet,* 13**,** 69-78.

IMURA, A., IWANO, A., TOHYAMA, O., TSUJI, Y., NOZAKI, K., HASHIMOTO, N., FUJIMORI, T. & NABESHIMA, Y. 2004. Secreted Klotho protein in sera and CSF: implication for post-translational cleavage in release of Klotho protein from cell membrane. *FEBS Lett,* 565**,** 143-7.

INAGAKI, T., CHOI, M., MOSCHETTA, A., PENG, L., CUMMINS, C. L., MCDONALD, J. G., LUO, G., JONES, S. A., GOODWIN, B., RICHARDSON, J. A., GERARD, R. D., REPA, J. J., MANGELSDORF, D. J. & KLIEWER, S. A. 2005. Fibroblast growth factor 15 functions as an enterohepatic signal to regulate bile acid homeostasis. *Cell Metab,* 2**,** 217-25.

INAGAKI, T., DUTCHAK, P., ZHAO, G., DING, X., GAUTRON, L., PARAMESWARA, V., LI, Y., GOETZ, R., MOHAMMADI, M., ESSER, V., ELMQUIST, J. K., GERARD, R. D., BURGESS, S. C., HAMMER, R. E., MANGELSDORF, D. J. & KLIEWER, S. A. 2007. Endocrine regulation of the fasting response by PPARalpha-mediated induction of fibroblast growth factor 21. *Cell Metab,* 5**,** 415-25.

INTERNATIONAL CANCER GENOME, C., HUDSON, T. J., ANDERSON, W., ARTEZ, A., BARKER, A. D., BELL, C., BERNABE, R. R., BHAN, M. K., CALVO, F., EEROLA, I., GERHARD, D. S., GUTTMACHER, A., GUYER, M., HEMSLEY, F. M., JENNINGS, J. L., KERR, D., KLATT, P., KOLAR, P., KUSADA, J., LANE, D. P., LAPLACE, F., YOUYONG, L., NETTEKOVEN, G., OZENBERGER, B., PETERSON, J., RAO, T. S., REMACLE, J., SCHAFER, A. J., SHIBATA, T., STRATTON, M. R., VOCKLEY, J. G., WATANABE, K., YANG, H., YUEN, M. M., KNOPPERS, B. M., BOBROW, M., CAMBON-THOMSEN, A., DRESSLER, L. G., DYKE, S. O., JOLY, Y., KATO, K., KENNEDY, K. L., NICOLAS, P., PARKER, M. J., RIAL-SEBBAG, E., ROMEO-CASABONA, C. M., SHAW, K. M., WALLACE, S., WIESNER, G. L., ZEPS, N., LICHTER, P., BIANKIN, A. V., CHABANNON, C., CHIN, L., CLEMENT, B., DE ALAVA, E., DEGOS, F., FERGUSON, M. L., GEARY, P., HAYES, D. N., HUDSON, T. J., JOHNS, A. L., KASPRZYK, A., NAKAGAWA, H., PENNY, R., PIRIS, M. A., SARIN, R., SCARPA, A., SHIBATA, T., VAN DE VIJVER, M., FUTREAL, P. A., ABURATANI, H., BAYES, M., BOTWELL, D. D., CAMPBELL, P. J., ESTIVILL, X., GERHARD, D. S., GRIMMOND, S. M., GUT, I., HIRST, M., LOPEZ-OTIN, C., MAJUMDER, P., MARRA, M., MCPHERSON, J. D., NAKAGAWA, H., NING, Z., PUENTE, X. S., RUAN, Y., SHIBATA, T., STRATTON, M. R., STUNNENBERG, H. G., SWERDLOW, H., VELCULESCU, V. E., WILSON, R. K., XUE, H. H., YANG, L., SPELLMAN, P. T., BADER, G. D., BOUTROS, P. C., CAMPBELL, P. J., et al. 2010. International network of cancer genome projects. *Nature,* 464**,** 993-8.

ITO, S., FUJIMORI, T., HAYASHIZAKI, Y. & NABESHIMA, Y. 2002. Identification of a novel mouse membrane-bound family 1 glycosidase-like protein, which carries an atypical active site structure. *Biochim Biophys Acta,* 1576**,** 341-5.

ITO, S., KINOSHITA, S., SHIRAISHI, N., NAKAGAWA, S., SEKINE, S., FUJIMORI, T. & NABESHIMA, Y. I. 2000. Molecular cloning and expression analyses of mouse betaklotho, which encodes a novel Klotho family protein. *Mech Dev,* 98**,** 115-9.

ITOH, N., OHTA, H. & KONISHI, M. 2015. Endocrine FGFs: Evolution, Physiology, Pathophysiology, and Pharmacotherapy. *Front Endocrinol (Lausanne),* 6**,** 154.

ITOH, N. & ORNITZ, D. M. 2004. Evolution of the Fgf and Fgfr gene families. *Trends Genet,* 20**,** 563-9.

ITOH, N. & ORNITZ, D. M. 2008. Functional evolutionary history of the mouse Fgf gene family. *Dev Dyn,* 237**,** 18-27.
ITOH, N. & ORNITZ, D. M. 2011. Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease. *J Biochem,* 149**,** 121-30.
JABS, E. W., LI, X., SCOTT, A. F., MEYERS, G., CHEN, W., ECCLES, M., MAO, J. I., CHARNAS, L. R., JACKSON, C. E. & JAYE, M. 1994. Jackson-Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor 2. *Nat Genet,* 8**,** 275-9.
JOHNSON, D. E., LEE, P. L., LU, J. & WILLIAMS, L. T. 1990. Diverse forms of a receptor for acidic and basic fibroblast growth factors. *Mol Cell Biol,* 10**,** 4728-36.
JURA, N., ENDRES, N. F., ENGEL, K., DEINDL, S., DAS, R., LAMERS, M. H., WEMMER, D. E., ZHANG, X. & KURIYAN, J. 2009. Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. *Cell,* 137**,** 1293-307.
KAIBORI, M., SAKAI, K., ISHIZAKI, M., MATSUSHIMA, H., DE VELASCO, M. A., MATSUI, K., IIDA, H., KITADE, H., KWON, A. H., NAGANO, H., WADA, H., HAJI, S., TSUKAMOTO, T., KANAZAWA, A., TAKEDA, Y., TAKEMURA, S., KUBO, S. & NISHIO, K. 2016. Increased FGF19 copy number is frequently detected in hepatocellular carcinoma with a complete response after sorafenib treatment. *Oncotarget,* 7**,** 49091-49098.
KALININA, J., BYRON, S. A., MAKARENKOVA, H. P., OLSEN, S. K., ELISEENKOVA, A. V., LAROCHELLE, W. J., DHANABAL, M., BLAIS, S., ORNITZ, D. M., DAY, L. A., NEUBERT, T. A., POLLOCK, P. M. & MOHAMMADI, M. 2009. Homodimerization controls the fibroblast growth factor 9 subfamily's receptor binding and heparan sulfate-dependent diffusion in the extracellular matrix. *Mol Cell Biol,* 29**,** 4663-78.
KALININA, J., DUTTA, K., ILGHARI, D., BEENKEN, A., GOETZ, R., ELISEENKOVA, A. V., COWBURN, D. & MOHAMMADI, M. 2012. The alternatively spliced acid box region plays a key role in FGF receptor autoinhibition. *Structure,* 20**,** 77-88.
KANG, S., ELF, S., DONG, S., HITOSUGI, T., LYTHGOE, K., GUO, A., RUAN, H., LONIAL, S., KHOURY, H. J., WILLIAMS, I. R., LEE, B. H., ROESEL, J. L., KARSENTY, G., HANAUER, A., TAUNTON, J., BOGGON, T. J., GU, T. L. & CHEN, J. 2009. Fibroblast growth factor receptor 3 associates with and tyrosine phosphorylates p90 RSK2, leading to RSK2 activation that mediates hematopoietic transformation. *Mol Cell Biol,* 29**,** 2105-17.
KHARITONENKOV, A., BEALS, J. M., MICANOVIC, R., STRIFLER, B. A., RATHNACHALAM, R., WROBLEWSKI, V. J., LI, S., KOESTER, A., FORD, A. M., COSKUN, T., DUNBAR, J. D., CHENG, C. C., FRYE, C. C., BUMOL, T. F. & MOLLER, D. E. 2013. Rational design of a fibroblast growth factor 21-based clinical candidate, LY2405319. *PLoS One,* 8**,** e58575.
KHARITONENKOV, A., DUNBAR, J. D., BINA, H. A., BRIGHT, S., MOYERS, J. S., ZHANG, C., DING, L., MICANOVIC, R., MEHRBOD, S. F., KNIERMAN, M. D., HALE, J. E., COSKUN, T. & SHANAFELT, A. B. 2008. FGF-21/FGF-21 receptor interaction and activation is determined by beta Klotho. *Journal of Cellular Physiology,* 215**,** 1-7.
KHARITONENKOV, A., SHIYANOVA, T. L., KOESTER, A., FORD, A. M., MICANOVIC, R., GALBREATH, E. J., SANDUSKY, G. E., HAMMOND, L. J., MOYERS, J. S., OWENS, R. A., GROMADA, J., BROZINICK, J. T., HAWKINS, E. D., WROBLEWSKI, V. J., LI, D. S., MEHRBOD, F., JASKUNAS, S. R. & SHANAFELT, A. B. 2005. FGF-21 as a novel metabolic regulator. *J Clin Invest,* 115**,** 1627-35.
KILKENNY, D. M. & ROCHELEAU, J. V. 2016. The FGF21 Receptor Signaling Complex: Klothobeta, FGFR1c, and Other Regulatory Interactions. *Vitam Horm,* 101**,** 17-58.
KIM, B., WANG, S., LEE, J. M., JEONG, Y., AHN, T., SON, D. S., PARK, H. W., YOO, H. S., SONG, Y. J., LEE, E., OH, Y. M., LEE, S. B., CHOI, J., MURRAY, J. C., ZHOU, Y., SONG, P. H., KIM, K. A. & WEINER, L. M. 2015. Synthetic lethal screening reveals FGFR as one of the combinatorial targets to overcome resistance to Met-targeted therapy. *Oncogene,* 34**,** 1083-93.
KIM, S. M., KIM, H., YUN, M. R., KANG, H. N., PYO, K. H., PARK, H. J., LEE, J. M., CHOI, H. M., ELLINGHAUS, P., OCKER, M., PAIK, S., KIM, H. R. & CHO, B. C. 2016. Activation of the Met kinase confers acquired drug resistance in FGFR-targeted lung cancer therapy. *Oncogenesis,* 5**,** e241.
KIMURA, T., SUZUKI, H., OHASHI, T., ASANO, K., KIYOTA, H. & ETO, Y. 2001. The incidence of thanatophoric dysplasia mutations in FGFR3 gene is higher in low-grade or superficial bladder carcinomas. *Cancer,* 92**,** 2555-61.
KIR, S., KLIEWER, S. A. & MANGELSDORF, D. J. 2011. Roles of FGF19 in liver metabolism. *Cold Spring Harb Symp Quant Biol,* 76**,** 139-44.

KITAYAMA, H., KANAKURA, Y., FURITSU, T., TSUJIMURA, T., ORITANI, K., IKEDA, H., SUGAHARA, H., MITSUI, H., KANAYAMA, Y., KITAMURA, Y. & ET AL. 1995. Constitutively activating mutations of c-kit receptor tyrosine kinase confer factor-independent growth and tumorigenicity of factor-dependent hematopoietic cell lines. *Blood,* 85**,** 790-8.

KOBASHIGAWA, Y., AMANO, S., YOKOGAWA, M., KUMETA, H., MORIOKA, H., INOUYE, M., SCHLESSINGER, J. & INAGAKI, F. 2015. Structural analysis of the mechanism of phosphorylation of a critical autoregulatory tyrosine residue in FGFR1 kinase domain. *Genes Cells,* 20**,** 860-70.

KOBAYASHI, S., BOGGON, T. J., DAYARAM, T., JANNE, P. A., KOCHER, O., MEYERSON, M., JOHNSON, B. E., ECK, M. J., TENEN, D. G. & HALMOS, B. 2005. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med,* 352**,** 786-92.

KOLUMAM, G., CHEN, M. Z., TONG, R., ZAVALA-SOLORIO, J., KATES, L., VAN BRUGGEN, N., ROSS, J., WYATT, S. K., GANDHAM, V. D., CARANO, R. A., DUNSHEE, D. R., WU, A. L., HALEY, B., ANDERSON, K., WARMING, S., RAIRDAN, X. Y., LEWIN-KOH, N., ZHANG, Y., GUTIERREZ, J., BARUCH, A., GELZLEICHTER, T. R., STEVENS, D., RAJAN, S., BAINBRIDGE, T. W., VERNES, J. M., MENG, Y. G., ZIAI, J., SORIANO, R. H., BRAUER, M. J., CHEN, Y., STAWICKI, S., KIM, H. S., COMPS-AGRAR, L., LUIS, E., SPIESS, C., WU, Y., ERNST, J. A., MCGUINNESS, O. P., PETERSON, A. S. & SONODA, J. 2015. Sustained Brown Fat Stimulation and Insulin Sensitization by a Humanized Bispecific Antibody Agonist for Fibroblast Growth Factor Receptor 1/betaKlotho Complex. *EBioMedicine,* 2**,** 730-43.

KONG, B., HUANG, J., ZHU, Y., LI, G., WILLIAMS, J., SHEN, S., ALEKSUNES, L. M., RICHARDSON, J. R., APTE, U., RUDNICK, D. A. & GUO, G. L. 2014. Fibroblast growth factor 15 deficiency impairs liver regeneration in mice. *Am J Physiol Gastrointest Liver Physiol,* 306**,** G893-902.

KONG, M., WANG, C. S. & DONOGHUE, D. J. 2002. Interaction of fibroblast growth factor receptor 3 and the adapter protein SH2-B. A role in STAT5 activation. *J Biol Chem,* 277**,** 15962-70.

KORNEV, A. P., HASTE, N. M., TAYLOR, S. S. & EYCK, L. F. 2006. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A,* 103**,** 17783-8.

KORNEV, A. P., TAYLOR, S. S. & TEN EYCK, L. F. 2008. A helix scaffold for the assembly of active protein kinases. *Proc Natl Acad Sci U S A,* 105**,** 14377-82.

KOUHARA, H., HADARI, Y. R., SPIVAK-KROIZMAN, T., SCHILLING, J., BAR-SAGI, D., LAX, I. & SCHLESSINGER, J. 1997. A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell,* 89**,** 693-702.

KOVESDY, C. P. & QUARLES, L. D. 2016. FGF23 from bench to bedside. *Am J Physiol Renal Physiol,* 310**,** F1168-74.

KRAUSE, D. S. & VAN ETTEN, R. A. 2005. Tyrosine kinases as targets for cancer therapy. *N Engl J Med,* 353**,** 172-87.

KREJCI, P., SALAZAR, L., KASHIWADA, T. A., CHLEBOVA, K., SALASOVA, A., THOMPSON, L. M., BRYJA, V., KOZUBIK, A. & WILCOX, W. R. 2008. Analysis of STAT1 activation by six FGFR3 mutants associated with skeletal dysplasia undermines dominant role of STAT1 in FGFR3 signaling in cartilage. *PLoS One,* 3**,** e3961.

KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol,* 305**,** 567-80.

KURO-O, M., MATSUMURA, Y., AIZAWA, H., KAWAGUCHI, H., SUGA, T., UTSUGI, T., OHYAMA, Y., KURABAYASHI, M., KANAME, T., KUME, E., IWASAKI, H., IIDA, A., SHIRAKI-IIDA, T., NISHIKAWA, S., NAGAI, R. & NABESHIMA, Y. I. 1997. Mutation of the mouse klotho gene leads to a syndrome resembling ageing. *Nature,* 390**,** 45-51.

KUROSU, H., CHOI, M., OGAWA, Y., DICKSON, A. S., GOETZ, R., ELISEENKOVA, A. V., MOHAMMADI, M., ROSENBLATT, K. P., KLIEWER, S. A. & KURO-O, M. 2007. Tissue-specific expression of betaKlotho and fibroblast growth factor (FGF) receptor isoforms determines metabolic activity of FGF19 and FGF21. *J Biol Chem,* 282**,** 26687-95.

KUROSU, H. & KURO, O. M. 2009. The Klotho gene family as a regulator of endocrine fibroblast growth factors. *Mol Cell Endocrinol,* 299**,** 72-8.

KUROSU, H., OGAWA, Y., MIYOSHI, M., YAMAMOTO, M., NANDI, A., ROSENBLATT, K. P., BAUM, M. G., SCHIAVI, S., HU, M. C., MOE, O. W. & KURO-O, M. 2006. Regulation of fibroblast growth factor-23 signaling by klotho. *J Biol Chem,* 281**,** 6120-3.

KUROSU, H., YAMAMOTO, M., CLARK, J. D., PASTOR, J. V., NANDI, A., GURNANI, P., MCGUINNESS, O. P., CHIKUDA, H., YAMAGUCHI, M., KAWAGUCHI, H., SHIMOMURA, I., TAKAYAMA, Y., HERZ, J., KAHN, C. R., ROSENBLATT, K. P. & KURO-O, M. 2005. Suppression of aging in mice by the hormone Klotho. *Science,* 309**,** 1829-33.

KWAK, Y., CHO, H., HUR, W. & SIM, T. 2015. Antitumor Effects and Mechanisms of AZD4547 on FGFR2-Deregulated Endometrial Cancer Cells. *Mol Cancer Ther,* 14**,** 2292-302.

LAEDERICH, M. B., DEGNIN, C. R., LUNSTRUM, G. P., HOLDEN, P. & HORTON, W. A. 2011. Fibroblast growth factor receptor 3 (FGFR3) is a strong heat shock protein 90 (Hsp90) client: implications for therapeutic manipulation. *J Biol Chem,* 286**,** 19597-604.

LEE, J., JEONG, D. J., KIM, J., LEE, S., PARK, J. H., CHANG, B., JUNG, S. I., YI, L., HAN, Y., YANG, Y., KIM, K. I., LIM, J. S., YANG, I., JEON, S., BAE, D. H., KIM, C. J. & LEE, M. S. 2010. The anti-aging gene KLOTHO is a novel target for epigenetic silencing in human cervical carcinoma. *Mol Cancer,* 9**,** 109.

LEE, J. H., GIANNIKOPOULOS, P., DUNCAN, S. A., WANG, J., JOHANSEN, C. T., BROWN, J. D., PLUTZKY, J., HEGELE, R. A., GLIMCHER, L. H. & LEE, A. H. 2011. The transcription factor cyclic AMP-responsive element-binding protein H regulates triglyceride metabolism. *Nat Med,* 17**,** 812-5.

LEES, J. G., MILES, A. J., WIEN, F. & WALLACE, B. A. 2006. A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics,* 22**,** 1955-62.

LEES, J. G., SMITH, B. R., WIEN, F., MILES, A. J. & WALLACE, B. A. 2004. CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal Biochem,* 332**,** 285-9.

LEMMON, M. A. & SCHLESSINGER, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell,* 141**,** 1117-34.

LEW, E. D., FURDUI, C. M., ANDERSON, K. S. & SCHLESSINGER, J. 2009. The precise sequence of FGF receptor autophosphorylation is kinetically driven and is disrupted by oncogenic mutations. *Sci Signal,* 2**,** ra6.

LIAO, R. G., JUNG, J., TCHAICHA, J., WILKERSON, M. D., SIVACHENKO, A., BEAUCHAMP, E. M., LIU, Q., PUGH, T. J., PEDAMALLU, C. S., HAYES, D. N., GRAY, N. S., GETZ, G., WONG, K. K., HADDAD, R. I., MEYERSON, M. & HAMMERMAN, P. S. 2013. Inhibitor-sensitive FGFR2 and FGFR3 mutations in lung squamous cell carcinoma. *Cancer Res,* 73**,** 5195-205.

LIEVENS, P. M. & LIBOI, E. 2003. The thanatophoric dysplasia type II mutation hampers complete maturation of fibroblast growth factor receptor 3 (FGFR3), which activates signal transducer and activator of transcription 1 (STAT1) from the endoplasmic reticulum. *J Biol Chem,* 278**,** 17344-9.

LIM, K., GROEN, A., MOLOSTVOV, G., LU, T., LILLEY, K. S., SNEAD, D., JAMES, S., WILKINSON, I. B., TING, S., HSIAO, L. L., HIEMSTRA, T. F. & ZEHNDER, D. 2015. alpha-Klotho Expression in Human Tissues. *J Clin Endocrinol Metab,* 100**,** E1308-18.

LIN, B. C., WANG, M., BLACKMORE, C. & DESNOYERS, L. R. 2007. Liver-specific activities of FGF19 require Klotho beta. *J Biol Chem,* 282**,** 27277-84.

LIN, C. C., MELO, F. A., GHOSH, R., SUEN, K. M., STAGG, L. J., KIRKPATRICK, J., AROLD, S. T., AHMED, Z. & LADBURY, J. E. 2012. Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell,* 149**,** 1514-24.

LIU, S., GUO, R., SIMPSON, L. G., XIAO, Z. S., BURNHAM, C. E. & QUARLES, L. D. 2003. Regulation of fibroblastic growth factor 23 expression but not degradation by PHEX. *J Biol Chem,* 278**,** 37419-26.

LIU, Y., MA, J., BEENKEN, A., SRINIVASAN, L., ELISEENKOVA, A. V. & MOHAMMADI, M. 2017. Regulation of Receptor Binding Specificity of FGF9 by an Autoinhibitory Homodimerization. *Structure*.

LOBB, R. R. & FETT, J. W. 1984. Purification of two distinct growth factors from bovine neural tissue by heparin affinity chromatography. *Biochemistry,* 23**,** 6295-9.

LUNDASEN, T., HUNT, M. C., NILSSON, L. M., SANYAL, S., ANGELIN, B., ALEXSON, S. E. & RUDLING, M. 2007. PPARalpha is a key regulator of hepatic FGF21. *Biochem Biophys Res Commun,* 360**,** 437-40.

LUO, Y., YANG, C., LU, W., XIE, R., JIN, C., HUANG, P., WANG, F. & MCKEEHAN, W. L. 2010. Metabolic regulator betaKlotho interacts with fibroblast growth factor receptor 4 (FGFR4) to induce apoptosis and inhibit tumor cell proliferation. *J Biol Chem,* 285**,** 30069-78.

MAKARENKOVA, H. P., HOFFMAN, M. P., BEENKEN, A., ELISEENKOVA, A. V., MEECH, R., TSAU, C., PATEL, V. N., LANG, R. A. & MOHAMMADI, M. 2009. Differential interactions of FGFs with heparan sulfate control gradient formation and branching morphogenesis. *Sci Signal,* 2**,** ra55.
MANAVALAN, P. & JOHNSON, W. C., JR. 1987. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem,* 167**,** 76-85.
MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-34.
MAO, D., WACHTER, E. & WALLACE, B. A. 1982. Folding of the mitochondrial proton adenosinetriphosphatase proteolipid channel in phospholipid vesicles. *Biochemistry,* 21**,** 4960-8.
MARKAN, K. R., NABER, M. C., AMEKA, M. K., ANDEREGG, M. D., MANGELSDORF, D. J., KLIEWER, S. A., MOHAMMADI, M. & POTTHOFF, M. J. 2014. Circulating FGF21 is liver derived and enhances glucose uptake during refeeding and overfeeding. *Diabetes,* 63**,** 4057-63.
MATSUMURA, Y., AIZAWA, H., SHIRAKI-IIDA, T., NAGAI, R., KURO-O, M. & NABESHIMA, Y. 1998. Identification of the human klotho gene and its two transcripts encoding membrane and secreted klotho protein. *Biochem Biophys Res Commun,* 242**,** 626-30.
MCDONELL, L. M., KERNOHAN, K. D., BOYCOTT, K. M. & SAWYER, S. L. 2015. Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. *Hum Mol Genet,* 24**,** R60-6.
MCSKIMMING, D. I., DASTGHEIB, S., TALEVICH, E., NARAYANAN, A., KATIYAR, S., TAYLOR, S. S., KOCHUT, K. & KANNAN, N. 2015. ProKinO: a unified resource for mining the cancer kinome. *Hum Mutat,* 36**,** 175-86.
MCWHIRTER, J. R., GOULDING, M., WEINER, J. A., CHUN, J. & MURRE, C. 1997. A novel fibroblast growth factor gene expressed in the developing nervous system is a downstream target of the chimeric homeodomain oncoprotein E2A-Pbx1. *Development,* 124**,** 3221-32.
MEHARENA, H. S., CHANG, P., KESHWANI, M. M., ORUGANTY, K., NENE, A. K., KANNAN, N., TAYLOR, S. S. & KORNEV, A. P. 2013. Deciphering the structural basis of eukaryotic protein kinase regulation. *PLoS Biol,* 11**,** e1001680.
MELLOR, H. R. 2014. Targeted inhibition of the FGF19-FGFR4 pathway in hepatocellular carcinoma; translational safety considerations. *Liver Int,* 34**,** e1-9.
MENCKE, R., OLAUSON, H. & HILLEBRANDS, J. L. 2017. Effects of Klotho on fibrosis and cancer: A renal focus on mechanisms and therapeutic strategies. *Adv Drug Deliv Rev*.
MICANOVIC, R., RACHES, D. W., DUNBAR, J. D., DRIVER, D. A., BINA, H. A., DICKINSON, C. D. & KHARITONENKOV, A. 2009. Different roles of N- and C- termini in the functional activity of FGF21. *J Cell Physiol,* 219**,** 227-34.
MICSONAI, A., WIEN, F., KERNYA, L., LEE, Y. H., GOTO, Y., REFREGIERS, M. & KARDOS, J. 2015. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci U S A,* 112**,** E3095-103.
MIKI, T., BOTTARO, D. P., FLEMING, T. P., SMITH, C. L., BURGESS, W. H., CHAN, A. M. & AARONSON, S. A. 1992. Determination of ligand-binding specificity by alternative splicing: two distinct growth factor receptors encoded by a single gene. *Proc Natl Acad Sci U S A,* 89**,** 246-50.
MILBURN, D., LASKOWSKI, R. A. & THORNTON, J. M. 1998. Sequences annotated by structure: a tool to facilitate the use of structural information in sequence analysis. *Protein Eng,* 11**,** 855-9.
MING, A. Y., YOO, E., VORONTSOV, E. N., ALTAMENTOVA, S. M., KILKENNY, D. M. & ROCHELEAU, J. V. 2012. Dynamics and Distribution of Klothobeta (KLB) and fibroblast growth factor receptor-1 (FGFR1) in living cells reveal the fibroblast growth factor-21 (FGF21)-induced receptor complex. *J Biol Chem,* 287**,** 19997-20006.
MIYAKE, M., ISHII, M., KOYAMA, N., KAWASHIMA, K., KODAMA, T., ANAI, S., FUJIMOTO, K., HIRAO, Y. & SUGANO, K. 2010. 1-tert-butyl-3-[6-(3,5-dimethoxy-phenyl)-2-(4-diethylamino-butylamino)-pyrido[2,3 -d]pyrimidin-7-yl]-urea (PD173074), a selective tyrosine kinase inhibitor of fibroblast growth factor receptor-3 (FGFR3), inhibits cell proliferation of bladder cancer carrying the FGFR3 gene mutation along with up-regulation of p27/Kip1 and G1/G0 arrest. *J Pharmacol Exp Ther,* 332**,** 795-802.
MIZUNO, I., TAKAHASHI, Y., OKIMURA, Y., KAJI, H. & CHIHARA, K. 2001. Upregulation of the klotho gene expression by thyroid hormone and during adipose differentiation in 3T3-L1 adipocytes. *Life Sci,* 68**,** 2917-23.

MOHAMMADI, M., DIKIC, I., SOROKIN, A., BURGESS, W. H., JAYE, M. & SCHLESSINGER, J. 1996a. Identification of six novel autophosphorylation sites on fibroblast growth factor receptor 1 and elucidation of their importance in receptor activation and signal transduction. *Mol Cell Biol,* 16**,** 977-89.

MOHAMMADI, M., HONEGGER, A. M., ROTIN, D., FISCHER, R., BELLOT, F., LI, W., DIONNE, C. A., JAYE, M., RUBINSTEIN, M. & SCHLESSINGER, J. 1991. A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. *Mol Cell Biol,* 11**,** 5068-78.

MOHAMMADI, M., OLSEN, S. K. & GOETZ, R. 2005a. A protein canyon in the FGF-FGF receptor dimer selects from an a la carte menu of heparan sulfate motifs. *Curr Opin Struct Biol,* 15**,** 506-16.

MOHAMMADI, M., OLSEN, S. K. & IBRAHIMI, O. A. 2005b. Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev,* 16**,** 107-37.

MOHAMMADI, M., SCHLESSINGER, J. & HUBBARD, S. R. 1996b. Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell,* 86**,** 577-87.

MU, J., PINKSTAFF, J., LI, Z., SKIDMORE, L., LI, N., MYLER, H., DALLAS-YANG, Q., PUTNAM, A. M., YAO, J., BUSSELL, S., WU, M., NORMAN, T. C., RODRIGUEZ, C. G., KIMMEL, B., METZGER, J. M., MANIBUSAN, A., LEE, D., ZALLER, D. M., ZHANG, B. B., DIMARCHI, R. D., BERGER, J. P. & AXELROD, D. W. 2012. FGF21 analogs of sustained action enabled by orthogonal biosynthesis demonstrate enhanced antidiabetic pharmacology in rodents. *Diabetes,* 61**,** 505-12.

MUENKE, M., SCHELL, U., HEHR, A., ROBIN, N. H., LOSKEN, H. W., SCHINZEL, A., PULLEYN, L. J., RUTLAND, P., REARDON, W., MALCOLM, S. & ET AL. 1994. A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nat Genet,* 8**,** 269-74.

NAGATA, H., WOROBEC, A. S., OH, C. K., CHOWDHURY, B. A., TANNENBAUM, S., SUZUKI, Y. & METCALFE, D. D. 1995. Identification of a point mutation in the catalytic domain of the protooncogene c-kit in peripheral blood mononuclear cells of patients who have mastocytosis with an associated hematologic disorder. *Proc Natl Acad Sci U S A,* 92**,** 10560-4.

NAKANISHI, Y., AKIYAMA, N., TSUKAGUCHI, T., FUJII, T., SAKATA, K., SASE, H., ISOBE, T., MORIKAMI, K., SHINDOH, H., MIO, T., EBIIKE, H., TAKA, N., AOKI, Y. & ISHII, N. 2014. The fibroblast growth factor receptor genetic status as a potential predictor of the sensitivity to CH5183284/Debio 1347, a novel selective FGFR inhibitor. *Mol Cancer Ther,* 13**,** 2547-58.

NASKI, M. C., WANG, Q., XU, J. & ORNITZ, D. M. 1996. Graded activation of fibroblast growth factor receptor 3 by mutations causing achondroplasia and thanatophoric dysplasia. *Nat Genet,* 13**,** 233-7.

NICHOLES, K., GUILLET, S., TOMLINSON, E., HILLAN, K., WRIGHT, B., FRANTZ, G. D., PHAM, T. A., DILLARD-TELM, L., TSAI, S. P., STEPHAN, J. P., STINSON, J., STEWART, T. & FRENCH, D. M. 2002. A mouse model of hepatocellular carcinoma: ectopic expression of fibroblast growth factor 19 in skeletal muscle of transgenic mice. *Am J Pathol,* 160**,** 2295-307.

NISHIMURA, T., NAKATAKE, Y., KONISHI, M. & ITOH, N. 2000. Identification of a novel FGF, FGF-21, preferentially expressed in the liver. *Biochim Biophys Acta,* 1492**,** 203-6.

NISHIMURA, T., UTSUNOMIYA, Y., HOSHIKAWA, M., OHUCHI, H. & ITOH, N. 1999. Structure and expression of a novel human FGF, FGF-19, expressed in the fetal brain. *Biochim Biophys Acta,* 1444**,** 148-51.

NOGOVA, L., SEQUIST, L. V., PEREZ GARCIA, J. M., ANDRE, F., DELORD, J. P., HIDALGO, M., SCHELLENS, J. H., CASSIER, P. A., CAMIDGE, D. R., SCHULER, M., VAISHAMPAYAN, U., BURRIS, H., TIAN, G. G., CAMPONE, M., WAINBERG, Z. A., LIM, W. T., LORUSSO, P., SHAPIRO, G. I., PARKER, K., CHEN, X., CHOUDHURY, S., RINGEISEN, F., GRAUS-PORTA, D., PORTER, D., ISAACS, R., BUETTNER, R. & WOLF, J. 2017. Evaluation of BGJ398, a Fibroblast Growth Factor Receptor 1-3 Kinase Inhibitor, in Patients With Advanced Solid Tumors Harboring Genetic Alterations in Fibroblast Growth Factor Receptors: Results of a Global Phase I, Dose-Escalation and Dose-Expansion Study. *J Clin Oncol,* 35**,** 157-165.

OGAWA, Y., KUROSU, H., YAMAMOTO, M., NANDI, A., ROSENBLATT, K. P., GOETZ, R., ELISEENKOVA, A. V., MOHAMMADI, M. & KURO-O, M. 2007. BetaKlotho is required

for metabolic activity of fibroblast growth factor 21. *Proc Natl Acad Sci U S A,* 104**,** 7432-7.
OLAUSON, H., MENCKE, R., HILLEBRANDS, J. L. & LARSSON, T. E. 2017. Tissue expression and source of circulating alphaKlotho. *Bone,* 100**,** 19-35.
OLSEN, S. K., GARBI, M., ZAMPIERI, N., ELISEENKOVA, A. V., ORNITZ, D. M., GOLDFARB, M. & MOHAMMADI, M. 2003. Fibroblast growth factor (FGF) homologous factors share structural but not functional homology with FGFs. *J Biol Chem,* 278**,** 34226-36.
OLSEN, S. K., IBRAHIMI, O. A., RAUCCI, A., ZHANG, F., ELISEENKOVA, A. V., YAYON, A., BASILICO, C., LINHARDT, R. J., SCHLESSINGER, J. & MOHAMMADI, M. 2004. Insights into the molecular basis for fibroblast growth factor receptor autoinhibition and ligand-binding promiscuity. *Proc Natl Acad Sci U S A,* 101**,** 935-40.
OLSEN, S. K., LI, J. Y., BROMLEIGH, C., ELISEENKOVA, A. V., IBRAHIMI, O. A., LAO, Z., ZHANG, F., LINHARDT, R. J., JOYNER, A. L. & MOHAMMADI, M. 2006. Structural basis by which alternative splicing modulates the organizer activity of FGF8 in the brain. *Genes Dev,* 20**,** 185-98.
ONG, S. H., GUY, G. R., HADARI, Y. R., LAKS, S., GOTOH, N., SCHLESSINGER, J. & LAX, I. 2000. FRS2 proteins recruit intracellular signaling pathways by binding to diverse targets on fibroblast growth factor and nerve growth factor receptors. *Mol Cell Biol,* 20**,** 979-89.
OPALINSKI, L., SOKOLOWSKA-WEDZINA, A., SZCZEPARA, M., ZAKRZEWSKA, M. & OTLEWSKI, J. 2017. Antibody-induced dimerization of FGFR1 promotes receptor endocytosis independently of its kinase activity. *Sci Rep,* 7**,** 7121.
ORNITZ, D. M. & ITOH, N. 2015. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip Rev Dev Biol,* 4**,** 215-66.
ORNITZ, D. M. & LEDER, P. 1992. Ligand specificity and heparin dependence of fibroblast growth factor receptors 1 and 3. *J Biol Chem,* 267**,** 16305-11.
ORNITZ, D. M., XU, J., COLVIN, J. S., MCEWEN, D. G., MACARTHUR, C. A., COULIER, F., GAO, G. & GOLDFARB, M. 1996. Receptor specificity of the fibroblast growth factor family. *J Biol Chem,* 271**,** 15292-7.
ORNITZ, D. M., YAYON, A., FLANAGAN, J. G., SVAHN, C. M., LEVI, E. & LEDER, P. 1992. Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Mol Cell Biol,* 12**,** 240-7.
ORR-URTREGER, A., BEDFORD, M. T., BURAKOVA, T., ARMAN, E., ZIMMER, Y., YAYON, A., GIVOL, D. & LONAI, P. 1993. Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR2). *Dev Biol,* 158**,** 475-86.
OSTROVSKY, O., BERMAN, B., GALLAGHER, J., MULLOY, B., FERNIG, D. G., DELEHEDDE, M. & RON, D. 2002. Differential effects of heparin saccharides on the formation of specific fibroblast growth factor (FGF) and FGF receptor complexes. *J Biol Chem,* 277**,** 2444-53.
OWEN, B. M., MANGELSDORF, D. J. & KLIEWER, S. A. 2015. Tissue-specific actions of the metabolic hormones FGF15/19 and FGF21. *Trends Endocrinol Metab,* 26**,** 22-9.
PANNIER, S., MARTINOVIC, J., HEUERTZ, S., DELEZOIDE, A. L., MUNNICH, A., SCHIBLER, L., SERRE, V. & LEGEAI-MALLET, L. 2009. Thanatophoric dysplasia caused by double missense FGFR3 mutations. *Am J Med Genet A,* 149A**,** 1296-301.
PAO, W., MILLER, V. A., POLITI, K. A., RIELY, G. J., SOMWAR, R., ZAKOWSKI, M. F., KRIS, M. G. & VARMUS, H. 2005. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med,* 2**,** e73.
PATANI, H., BUNNEY, T. D., THIYAGARAJAN, N., NORMAN, R. A., OGG, D., BREED, J., ASHFORD, P., POTTERTON, A., EDWARDS, M., WILLIAMS, S. V., THOMSON, G. S., PANG, C. S., KNOWLES, M. A., BREEZE, A. L., ORENGO, C., PHILLIPS, C. & KATAN, M. 2016. Landscape of activating cancer mutations in FGFR kinases and their differential responses to inhibitors in clinical use. *Oncotarget,* 7**,** 24252-68.
PELLEGRINI, L., BURKE, D. F., VON DELFT, F., MULLOY, B. & BLUNDELL, T. L. 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature,* 407**,** 1029-34.
PENG, W. C., LIN, X. & TORRES, J. 2009. The strong dimerization of the transmembrane domain of the fibroblast growth factor receptor (FGFR) is modulated by C-terminal juxtamembrane residues. *Protein Sci,* 18**,** 450-9.
PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods,* 8**,** 785-6.

PIETRAS, K., PAHLER, J., BERGERS, G. & HANAHAN, D. 2008. Functions of paracrine PDGF signaling in the proangiogenic tumor stroma revealed by pharmacological targeting. *PLoS Med,* 5**,** e19.
PILLAY, V., ALLAF, L., WILDING, A. L., DONOGHUE, J. F., COURT, N. W., GREENALL, S. A., SCOTT, A. M. & JOHNS, T. G. 2009. The plasticity of oncogene addiction: implications for targeted therapies directed to receptor tyrosine kinases. *Neoplasia,* 11**,** 448-58, 2 p following 458.
PLOTNIKOV, A. N., ELISEENKOVA, A. V., IBRAHIMI, O. A., SHRIVER, Z., SASISEKHARAN, R., LEMMON, M. A. & MOHAMMADI, M. 2001. Crystal structure of fibroblast growth factor 9 reveals regions implicated in dimerization and autoinhibition. *J Biol Chem,* 276**,** 4322-9.
PLOTNIKOV, A. N., SCHLESSINGER, J., HUBBARD, S. R. & MOHAMMADI, M. 1999. Structural basis for FGF receptor dimerization and activation. *Cell,* 98**,** 641-50.
POTTHOFF, M. J., KLIEWER, S. A. & MANGELSDORF, D. J. 2012. Endocrine fibroblast growth factors 15/19 and 21: from feast to famine. *Genes Dev,* 26**,** 312-24.
POWELL, M. A., SILL, M. W., GOODFELLOW, P. J., BENBROOK, D. M., LANKES, H. A., LESLIE, K. K., JESKE, Y., MANNEL, R. S., SPILLMAN, M. A., LEE, P. S., HOFFMAN, J. S., MCMEEKIN, D. S. & POLLOCK, P. M. 2014. A phase II trial of brivanib in recurrent or persistent endometrial cancer: an NRG Oncology/Gynecologic Oncology Group Study. *Gynecol Oncol,* 135**,** 38-43.
POWNALL, M. E. & ISAACS, H. V. 2010. *FGF Signalling in Vertebrate Development.* San Rafael (CA).
QING, J., DU, X., CHEN, Y., CHAN, P., LI, H., WU, P., MARSTERS, S., STAWICKI, S., TIEN, J., TOTPAL, K., ROSS, S., STINSON, S., DORNAN, D., FRENCH, D., WANG, Q. R., STEPHAN, J. P., WU, Y., WIESMANN, C. & ASHKENAZI, A. 2009. Antibody-based targeting of FGFR3 in bladder carcinoma and t(4;14)-positive multiple myeloma in mice. *J Clin Invest,* 119**,** 1216-29.
QU, X., CARBE, C., TAO, C., POWERS, A., LAWRENCE, R., VAN KUPPEVELT, T. H., CARDOSO, W. V., GROBE, K., ESKO, J. D. & ZHANG, X. 2011. Lacrimal gland development and Fgf10-Fgfr2b signaling are controlled by 2-O- and 6-O-sulfated heparan sulfate. *J Biol Chem,* 286**,** 14435-44.
QU, X., PAN, Y., CARBE, C., POWERS, A., GROBE, K. & ZHANG, X. 2012. Glycosaminoglycan-dependent restriction of FGF diffusion is necessary for lacrimal gland development. *Development,* 139**,** 2730-9.
RAPRAEGER, A. C., KRUFKA, A. & OLWIN, B. B. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science,* 252**,** 1705-8.
REARDON, W., WINTER, R. M., RUTLAND, P., PULLEYN, L. J., JONES, B. M. & MALCOLM, S. 1994. Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nat Genet,* 8**,** 98-103.
RED BREWER, M., CHOI, S. H., ALVARADO, D., MORAVCEVIC, K., POZZI, A., LEMMON, M. A. & CARPENTER, G. 2009. The juxtamembrane region of the EGF receptor functions as an activation domain. *Mol Cell,* 34**,** 641-51.
ROBERTSON, S. C., TYNAN, J. & DONOGHUE, D. J. 2000. RTK mutations and human syndromes: when good receptors turn bad. *Trends Genet,* 16**,** 368.
ROGHANI, M. & MOSCATELLI, D. 2007. Prostate cells express two isoforms of fibroblast growth factor receptor 1 with different affinities for fibroblast growth factor-2. *Prostate,* 67**,** 115-24.
RONCHETTI, D., GRECO, A., COMPASSO, S., COLOMBO, G., DELL'ERA, P., OTSUKI, T., LOMBARDI, L. & NERI, A. 2001. Deregulated FGFR3 mutants in multiple myeloma cell lines with t(4;14): comparative analysis of Y373C, K650E and the novel G384D mutations. *Oncogene,* 20**,** 3553-62.
ROSTY, C., AUBRIOT, M. H., CAPPELLEN, D., BOURDIN, J., CARTIER, I., THIERY, J. P., SASTRE-GARAU, X. & RADVANYI, F. 2005. Clinical and biological characteristics of cervical neoplasias with FGFR3 mutation. *Mol Cancer,* 4**,** 15.
ROUSSEAU, F., BONAVENTURE, J., LEGEAI-MALLET, L., PELET, A., ROZET, J. M., MAROTEAUX, P., LE MERRER, M. & MUNNICH, A. 1994. Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature,* 371**,** 252-4.
RYU, E. K., CHO, K. J., KIM, J. K., HARMER, N. J., BLUNDELL, T. L. & KIM, K. H. 2006. Expression and purification of recombinant human fibroblast growth factor receptor in Escherichia coli. *Protein Expr Purif,* 49**,** 15-22.
SAKAMOTO, K., KAWATA, Y., MASUDA, Y., UMEMOTO, T., ITO, T., ASAMI, T., TAKEKAWA, S., OHTAKI, T. & INOOKA, H. 2016. Discovery of an artificial peptide agonist to the

fibroblast growth factor receptor 1c/betaKlotho complex from random peptide T7 phage display. *Biochem Biophys Res Commun,* 480**,** 55-60.
SAKSELA, O., MOSCATELLI, D., SOMMER, A. & RIFKIN, D. B. 1988. Endothelial cell-derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation. *J Cell Biol,* 107**,** 743-51.
SANDILANDS, E., AKBARZADEH, S., VECCHIONE, A., MCEWAN, D. G., FRAME, M. C. & HEATH, J. K. 2007. Src kinase modulates the activation, transport and signalling dynamics of fibroblast growth factor receptors. *EMBO Rep,* 8**,** 1162-9.
SANTOS-OCAMPO, S., COLVIN, J. S., CHELLAIAH, A. & ORNITZ, D. M. 1996. Expression and biological activity of mouse fibroblast growth factor-9. *J Biol Chem,* 271**,** 1726-31.
SARABIPOUR, S. 2017. Parallels and Distinctions in FGFR, VEGFR, and EGFR Mechanisms of Transmembrane Signaling. *Biochemistry,* 56**,** 3159-3173.
SARABIPOUR, S. & HRISTOVA, K. 2015. FGFR3 unliganded dimer stabilization by the juxtamembrane domain. *J Mol Biol,* 427**,** 1705-14.
SARABIPOUR, S. & HRISTOVA, K. 2016. Mechanism of FGF receptor dimerization and activation. *Nat Commun,* 7**,** 10262.
SAWEY, E. T., CHANRION, M., CAI, C., WU, G., ZHANG, J., ZENDER, L., ZHAO, A., BUSUTTIL, R. W., YEE, H., STEIN, L., FRENCH, D. M., FINN, R. S., LOWE, S. W. & POWERS, S. 2011. Identification of a therapeutic strategy targeting amplified FGF19 in liver cancer by Oncogenomic screening. *Cancer Cell,* 19**,** 347-58.
SCHLESSINGER, J., PLOTNIKOV, A. N., IBRAHIMI, O. A., ELISEENKOVA, A. V., YEH, B. K., YAYON, A., LINHARDT, R. J. & MOHAMMADI, M. 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell,* 6**,** 743-50.
SCHLESSINGER, J., SHECHTER, Y., CUATRECASAS, P., WILLINGHAM, M. C. & PASTAN, I. 1978. Quantitative determination of the lateral diffusion coefficients of the hormone-receptor complexes of insulin and epidermal growth factor on the plasma membrane of cultured fibroblasts. *Proc Natl Acad Sci U S A,* 75**,** 5353-7.
SCHMID, C., NEIDERT, M. C., TSCHOPP, O., SZE, L. & BERNAYS, R. L. 2013. Growth hormone and Klotho. *J Endocrinol,* 219**,** R37-57.
SESHAGIRI, S., STAWISKI, E. W., DURINCK, S., MODRUSAN, Z., STORM, E. E., CONBOY, C. B., CHAUDHURI, S., GUAN, Y., JANAKIRAMAN, V., JAISWAL, B. S., GUILLORY, J., HA, C., DIJKGRAAF, G. J., STINSON, J., GNAD, F., HUNTLEY, M. A., DEGENHARDT, J. D., HAVERTY, P. M., BOURGON, R., WANG, W., KOEPPEN, H., GENTLEMAN, R., STARR, T. K., ZHANG, Z., LARGAESPADA, D. A., WU, T. D. & DE SAUVAGE, F. J. 2012. Recurrent R-spondin fusions in colon cancer. *Nature,* 488**,** 660-4.
SHAH, D. R., SHAH, R. R. & MORGANROTH, J. 2013a. Tyrosine kinase inhibitors: their on-target toxicities as potential indicators of efficacy. *Drug Saf,* 36**,** 413-26.
SHAH, N. P., NICOLL, J. M., NAGAR, B., GORRE, M. E., PAQUETTE, R. L., KURIYAN, J. & SAWYERS, C. L. 2002. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell,* 2**,** 117-25.
SHAH, R. R., MORGANROTH, J. & SHAH, D. R. 2013b. Cardiovascular safety of tyrosine kinase inhibitors: with a special focus on cardiac repolarisation (QT interval). *Drug Saf,* 36**,** 295-316.
SHIANG, C. Y., QI, Y., WANG, B., LAZAR, V., WANG, J., FRASER SYMMANS, W., HORTOBAGYI, G. N., ANDRE, F. & PUSZTAI, L. 2010. Amplification of fibroblast growth factor receptor-1 in breast cancer and the effects of brivanib alaninate. *Breast Cancer Res Treat,* 123**,** 747-55.
SHIANG, R., THOMPSON, L. M., ZHU, Y. Z., CHURCH, D. M., FIELDER, T. J., BOCIAN, M., WINOKUR, S. T. & WASMUTH, J. J. 1994. Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell,* 78**,** 335-42.
SHIMADA, T., HASEGAWA, H., YAMAZAKI, Y., MUTO, T., HINO, R., TAKEUCHI, Y., FUJITA, T., NAKAHARA, K., FUKUMOTO, S. & YAMASHITA, T. 2004. FGF-23 is a potent regulator of vitamin D metabolism and phosphate homeostasis. *J Bone Miner Res,* 19**,** 429-35.
SHIMADA, T., MIZUTANI, S., MUTO, T., YONEYA, T., HINO, R., TAKEDA, S., TAKEUCHI, Y., FUJITA, T., FUKUMOTO, S. & YAMASHITA, T. 2001. Cloning and characterization of FGF23 as a causative factor of tumor-induced osteomalacia. *Proc Natl Acad Sci U S A,* 98**,** 6500-5.

SHIMADA, T., MUTO, T., URAKAWA, I., YONEYA, T., YAMAZAKI, Y., OKAWA, K., TAKEUCHI, Y., FUJITA, T., FUKUMOTO, S. & YAMASHITA, T. 2002. Mutant FGF-23 responsible for autosomal dominant hypophosphatemic rickets is resistant to proteolytic cleavage and causes hypophosphatemia in vivo. *Endocrinology,* 143**,** 3179-82.

SHIMIZU, A., TADA, K., SHUKUNAMI, C., HIRAKI, Y., KUROKAWA, T., MAGANE, N. & KUROKAWA-SEO, M. 2001. A novel alternatively spliced fibroblast growth factor receptor 3 isoform lacking the acid box domain is expressed during chondrogenic differentiation of ATDC5 cells. *J Biol Chem,* 276**,** 11031-40.

SHING, Y., FOLKMAN, J., SULLIVAN, R., BUTTERFIELD, C., MURRAY, J. & KLAGSBRUN, M. 1984. Heparin affinity: purification of a tumor-derived capillary endothelial cell growth factor. *Science,* 223**,** 1296-9.

SHIRAKI-IIDA, T., AIZAWA, H., MATSUMURA, Y., SEKINE, S., IIDA, A., ANAZAWA, H., NAGAI, R., KURO-O, M. & NABESHIMA, Y. 1998. Structure of the mouse klotho gene and its two transcripts encoding membrane and secreted protein. *FEBS Lett,* 424**,** 6-10.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

SINGH, D., CHAN, J. M., ZOPPOLI, P., NIOLA, F., SULLIVAN, R., CASTANO, A., LIU, E. M., REICHEL, J., PORRATI, P., PELLEGATTA, S., QIU, K., GAO, Z., CECCARELLI, M., RICCARDI, R., BRAT, D. J., GUHA, A., ALDAPE, K., GOLFINOS, J. G., ZAGZAG, D., MIKKELSEN, T., FINOCCHIARO, G., LASORELLA, A., RABADAN, R. & IAVARONE, A. 2012. Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science,* 337**,** 1231-5.

SINGLETON, K. R., KIM, J., HINZ, T. K., MAREK, L. A., CASAS-SELVES, M., HATHEWAY, C., TAN, A. C., DEGREGORI, J. & HEASLEY, L. E. 2013. A receptor tyrosine kinase network composed of fibroblast growth factor receptors, epidermal growth factor receptor, v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, and hepatocyte growth factor receptor drives growth and survival of head and neck squamous carcinoma cell lines. *Mol Pharmacol,* 83**,** 882-93.

SLEEMAN, M., FRASER, J., MCDONALD, M., YUAN, S., WHITE, D., GRANDISON, P., KUMBLE, K., WATSON, J. D. & MURISON, J. G. 2001. Identification of a new fibroblast growth factor receptor, FGFR5. *Gene,* 271**,** 171-82.

SMITH, R., DUGUAY, A., BAKKER, A., LI, P., WEISZMANN, J., THOMAS, M. R., ALBA, B. M., WU, X., GUPTE, J., YANG, L., STEVENS, J., HAMBURGER, A., SMITH, S., CHEN, J., KOMOROWSKI, R., MOORE, K. W., VENIANT, M. M. & LI, Y. 2013a. FGF21 can be mimicked in vitro and in vivo by a novel anti-FGFR1c/beta-Klotho bispecific protein. *PLoS One,* 8**,** e61432.

SMITH, R., DUGUAY, A., WEISZMANN, J., STANISLAUS, S., BELOUSKI, E., CAI, L., YIE, J., XU, J., GUPTE, J., WU, X. & LI, Y. 2013b. A novel approach to improve the function of FGF21. *BioDrugs,* 27**,** 159-66.

SOHL, C. D., RYAN, M. R., LUO, B., FREY, K. M. & ANDERSON, K. S. 2015. Illuminating the molecular mechanisms of tyrosine kinase inhibitor resistance for the FGFR1 gatekeeper mutation: the Achilles' heel of targeted therapy. *ACS Chem Biol,* 10**,** 1319-29.

SONG, Y., DING, J., JIN, R., JUNG, J., LI, S., YANG, J., WANG, A. & LI, Z. 2016. Expression and purification of FGF21 in Pichia pastoris and its effect on fibroblast-cell migration. *Mol Med Rep,* 13**,** 3619-26.

SORIA, J. C., DEBRAUD, F., BAHLEDA, R., ADAMO, B., ANDRE, F., DIENSTMANN, R., DELMONTE, A., CEREDA, R., ISAACSON, J., LITTEN, J., ALLEN, A., DUBOIS, F., SABA, C., ROBERT, R., D'INCALCI, M., ZUCCHETTI, M., CAMBONI, M. G. & TABERNERO, J. 2014. Phase I/IIa study evaluating the safety, efficacy, pharmacokinetics, and pharmacodynamics of lucitanib in advanced solid tumors. *Ann Oncol,* 25**,** 2244-51.

SORKIN, A. & VON ZASTROW, M. 2009. Endocytosis and signalling: intertwining molecular networks. *Nat Rev Mol Cell Biol,* 10**,** 609-22.

SPIELBERGER, R., STIFF, P., BENSINGER, W., GENTILE, T., WEISDORF, D., KEWALRAMANI, T., SHEA, T., YANOVICH, S., HANSEN, K., NOGA, S., MCCARTY, J., LEMAISTRE, C. F., SUNG, E. C., BLAZAR, B. R., ELHARDT, D., CHEN, M. G. & EMMANOUILIDES, C. 2004. Palifermin for oral mucositis after intensive therapy for hematologic cancers. *N Engl J Med,* 351**,** 2590-8.

STAUBER, D. J., DIGABRIELE, A. D. & HENDRICKSON, W. A. 2000. Structural interactions of fibroblast growth factor receptor with its ligands. *Proc Natl Acad Sci U S A,* 97**,** 49-54.

STOMMEL, J. M., KIMMELMAN, A. C., YING, H., NABIOULLIN, R., PONUGOTI, A. H., WIEDEMEYER, R., STEGH, A. H., BRADNER, J. E., LIGON, K. L., BRENNAN, C., CHIN, L. & DEPINHO, R. A. 2007. Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. *Science,* 318**,** 287-90.

STUHLMILLER, T. J., MILLER, S. M., ZAWISTOWSKI, J. S., NAKAMURA, K., BELTRAN, A. S., DUNCAN, J. S., ANGUS, S. P., COLLINS, K. A., GRANGER, D. A., REUTHER, R. A., GRAVES, L. M., GOMEZ, S. M., KUAN, P. F., PARKER, J. S., CHEN, X., SCIAKY, N., CAREY, L. A., EARP, H. S., JIN, J. & JOHNSON, G. L. 2015. Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep,* 11**,** 390-404.

SUZUKI, M., UEHARA, Y., MOTOMURA-MATSUZAKA, K., OKI, J., KOYAMA, Y., KIMURA, M., ASADA, M., KOMI-KURAMOCHI, A., OKA, S. & IMAMURA, T. 2008. betaKlotho is required for fibroblast growth factor (FGF) 21 signaling through FGF receptor (FGFR) 1c and FGFR3c. *Mol Endocrinol,* 22**,** 1006-14.

TAMBORINI, E., BONADIMAN, L., GRECO, A., ALBERTINI, V., NEGRI, T., GRONCHI, A., BERTULLI, R., COLECCHIA, M., CASALI, P. G., PIEROTTI, M. A. & PILOTTI, S. 2004. A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. *Gastroenterology,* 127**,** 294-9.

TAN, L., WANG, J., TANIZAKI, J., HUANG, Z., AREF, A. R., RUSAN, M., ZHU, S. J., ZHANG, Y., ERCAN, D., LIAO, R. G., CAPELLETTI, M., ZHOU, W., HUR, W., KIM, N., SIM, T., GAUDET, S., BARBIE, D. A., YEH, J. R., YUN, C. H., HAMMERMAN, P. S., MOHAMMADI, M., JANNE, P. A. & GRAY, N. S. 2014. Development of covalent inhibitors that can overcome resistance to first-generation FGFR kinase inhibitors. *Proc Natl Acad Sci U S A,* 111**,** E4869-77.

TANNER, Y. & GROSE, R. P. 2016. Dysregulated FGF signalling in neoplastic disorders. *Semin Cell Dev Biol,* 53**,** 126-35.

TAVORMINA, P. L., BELLUS, G. A., WEBSTER, M. K., BAMSHAD, M. J., FRALEY, A. E., MCINTOSH, I., SZABO, J., JIANG, W., JABS, E. W., WILCOX, W. R., WASMUTH, J. J., DONOGHUE, D. J., THOMPSON, L. M. & FRANCOMANO, C. A. 1999. A novel skeletal dysplasia with developmental delay and acanthosis nigricans is caused by a Lys650Met mutation in the fibroblast growth factor receptor 3 gene. *Am J Hum Genet,* 64**,** 722-31.

TAVORMINA, P. L., SHIANG, R., THOMPSON, L. M., ZHU, Y. Z., WILKIN, D. J., LACHMAN, R. S., WILCOX, W. R., RIMOIN, D. L., COHN, D. H. & WASMUTH, J. J. 1995. Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. *Nat Genet,* 9**,** 321-8.

TAYLOR, S. S. & KORNEV, A. P. 2011. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci,* 36**,** 65-77.

TERAI, H., SOEJIMA, K., YASUDA, H., NAKAYAMA, S., HAMAMOTO, J., ARAI, D., ISHIOKA, K., OHGINO, K., IKEMURA, S., SATO, T., YODA, S., SATOMI, R., NAOKI, K. & BETSUYAKU, T. 2013. Activation of the FGF2-FGFR1 autocrine pathway: a novel mechanism of acquired resistance to gefitinib in NSCLC. *Mol Cancer Res,* 11**,** 759-67.

THE UNIPROT, C. 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Res,* 45**,** D158-D169.

THORNTON, S. C., MUELLER, S. N. & LEVINE, E. M. 1983. Human endothelial cells: use of heparin in cloning and long-term serial cultivation. *Science,* 222**,** 623-5.

TOHYAMA, O., IMURA, A., IWANO, A., FREUND, J. N., HENRISSAT, B., FUJIMORI, T. & NABESHIMA, Y. 2004. Klotho is a novel beta-glucuronidase capable of hydrolyzing steroid beta-glucuronides. *J Biol Chem,* 279**,** 9777-84.

TOMLINSON, D. C., BALDO, O., HARNDEN, P. & KNOWLES, M. A. 2007a. FGFR3 protein expression and its relationship to mutation status and prognostic variables in bladder cancer. *J Pathol,* 213**,** 91-8.

TOMLINSON, D. C., HURST, C. D. & KNOWLES, M. A. 2007b. Knockdown by shRNA identifies S249C mutant FGFR3 as a potential therapeutic target in bladder cancer. *Oncogene,* 26**,** 5889-99.

TOMLINSON, D. C., L'HOTE, C. G., KENNEDY, W., PITT, E. & KNOWLES, M. A. 2005. Alternative splicing of fibroblast growth factor receptor 3 produces a secreted isoform that inhibits fibroblast growth factor-induced proliferation and is repressed in urothelial carcinoma cell lines. *Cancer Res,* 65**,** 10441-9.

TOMLINSON, E., FU, L., JOHN, L., HULTGREN, B., HUANG, X., RENZ, M., STEPHAN, J. P., TSAI, S. P., POWELL-BRAXTON, L., FRENCH, D. & STEWART, T. A. 2002.

Transgenic mice expressing human fibroblast growth factor-19 display increased metabolic rate and decreased adiposity. *Endocrinology,* 143**,** 1741-7.
TOUAT, M., ILEANA, E., POSTEL-VINAY, S., ANDRE, F. & SORIA, J. C. 2015. Targeting FGFR Signaling in Cancer. *Clin Cancer Res,* 21**,** 2684-94.
TRUDEL, S., ELY, S., FAROOQI, Y., AFFER, M., ROBBIANI, D. F., CHESI, M. & BERGSAGEL, P. L. 2004. Inhibition of fibroblast growth factor receptor 3 induces differentiation and apoptosis in t(4;14) myeloma. *Blood,* 103**,** 3521-8.
TRUEB, B. 2011. Biology of FGFRL1, the fifth fibroblast growth factor receptor. *Cell Mol Life Sci,* 68**,** 951-64.
TRUEB, B., ZHUANG, L., TAESCHLER, S. & WIEDEMANN, M. 2003. Characterization of FGFRL1, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues. *J Biol Chem,* 278**,** 33857-65.
TURNER, N. & GROSE, R. 2010. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer,* 10**,** 116-29.
UEMATSU, F., JANG, J. H., KAN, M., WANG, F., LUO, Y. & MCKEEHAN, W. L. 2001. Evidence that the intracellular domain of FGF receptor 2IIIb affects contact of the ectodomain with two FGF7 ligands. *Biochem Biophys Res Commun,* 283**,** 791-7.
URAKAWA, I., YAMAZAKI, Y., SHIMADA, T., IIJIMA, K., HASEGAWA, H., OKAWA, K., FUJITA, T., FUKUMOTO, S. & YAMASHITA, T. 2006. Klotho converts canonical FGF receptor into a specific receptor for FGF23. *Nature,* 444**,** 770-4.
URIARTE, I., FERNANDEZ-BARRENA, M. G., MONTE, M. J., LATASA, M. U., CHANG, H. C., CAROTTI, S., VESPASIANI-GENTILUCCI, U., MORINI, S., VICENTE, E., CONCEPCION, A. R., MEDINA, J. F., MARIN, J. J., BERASAIN, C., PRIETO, J. & AVILA, M. A. 2013. Identification of fibroblast growth factor 15 as a novel mediator of liver regeneration and its application in the prevention of post-resection liver failure in mice. *Gut,* 62**,** 899-910.
URIARTE, I., LATASA, M. U., CAROTTI, S., FERNANDEZ-BARRENA, M. G., GARCIA-IRIGOYEN, O., ELIZALDE, M., URTASUN, R., VESPASIANI-GENTILUCCI, U., MORINI, S., DE MINGO, A., MARI, M., CORRALES, F. J., PRIETO, J., BERASAIN, C. & AVILA, M. A. 2015. Ileal FGF15 contributes to fibrosis-associated hepatocellular carcinoma development. *Int J Cancer,* 136**,** 2469-75.
VALIEV, M., KAWAI, R., ADAMS, J. A. & WEARE, J. H. 2003. The role of the putative catalytic base in the phosphoryl transfer reaction in a protein kinase: first-principles calculations. *J Am Chem Soc,* 125**,** 9926-7.
VAN RHIJN, B. W., LURKIN, I., RADVANYI, F., KIRKELS, W. J., VAN DER KWAST, T. H. & ZWARTHOFF, E. C. 2001. The fibroblast growth factor receptor 3 (FGFR3) mutation is a strong indicator of superficial bladder cancer with low recurrence rate. *Cancer Res,* 61**,** 1265-8.
WANG, F., KAN, M., YAN, G., XU, J. & MCKEEHAN, W. L. 1995. Alternately spliced NH2-terminal immunoglobulin-like Loop I in the ectodomain of the fibroblast growth factor (FGF) receptor 1 lowers affinity for both heparin and FGF-1. *J Biol Chem,* 270**,** 10231-5.
WARE, K. E., MARSHALL, M. E., HEASLEY, L. R., MAREK, L., HINZ, T. K., HERCULE, P., HELFRICH, B. A., DOEBELE, R. C. & HEASLEY, L. E. 2010. Rapidly acquired resistance to EGFR tyrosine kinase inhibitors in NSCLC cell lines through de-repression of FGFR2 and FGFR3 expression. *PLoS One,* 5**,** e14117.
WEBSTER, M. K., D'AVIS, P. Y., ROBERTSON, S. C. & DONOGHUE, D. J. 1996. Profound ligand-independent kinase activation of fibroblast growth factor receptor 3 by the activation loop mutation responsible for a lethal skeletal dysplasia, thanatophoric dysplasia type II. *Mol Cell Biol,* 16**,** 4081-7.
WEISS, J., SOS, M. L., SEIDEL, D., PEIFER, M., ZANDER, T., HEUCKMANN, J. M., ULLRICH, R. T., MENON, R., MAIER, S., SOLTERMANN, A., MOCH, H., WAGENER, P., FISCHER, F., HEYNCK, S., KOKER, M., SCHOTTLE, J., LEENDERS, F., GABLER, F., DABOW, I., QUERINGS, S., HEUKAMP, L. C., BALKE-WANT, H., ANSEN, S., RAUH, D., BAESSMANN, I., ALTMULLER, J., WAINER, Z., CONRON, M., WRIGHT, G., RUSSELL, P., SOLOMON, B., BRAMBILLA, E., BRAMBILLA, C., LORIMIER, P., SOLLBERG, S., BRUSTUGUN, O. T., ENGEL-RIEDEL, W., LUDWIG, C., PETERSEN, I., SANGER, J., CLEMENT, J., GROEN, H., TIMENS, W., SIETSMA, H., THUNNISSEN, E., SMIT, E., HEIDEMAN, D., CAPPUZZO, F., LIGORIO, C., DAMIANI, S., HALLEK, M., BEROUKHIM, R., PAO, W., KLEBL, B., BAUMANN, M., BUETTNER, R., ERNESTUS, K., STOELBEN, E., WOLF, J., NURNBERG, P., PERNER, S. & THOMAS, R. K. 2010. Frequent and focal FGFR1 amplification associates with

therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci Transl Med,* 2**,** 62ra93.
WERNER, S., DUAN, D. S., DE VRIES, C., PETERS, K. G., JOHNSON, D. E. & WILLIAMS, L. T. 1992. Differential splicing in the extracellular region of fibroblast growth factor receptor 1 generates receptor variants with different ligand-binding specificities. *Mol Cell Biol,* 12**,** 82-8.
WHITE, K. E., CARN, G., LORENZ-DEPIEREUX, B., BENET-PAGES, A., STROM, T. M. & ECONS, M. J. 2001. Autosomal-dominant hypophosphatemic rickets (ADHR) mutations stabilize FGF-23. *Kidney Int,* 60**,** 2079-86.
WIEDEMANN, M. & TRUEB, B. 2000. Characterization of a novel protein (FGFRL1) from human cartilage related to FGF receptors. *Genomics,* 69**,** 275-9.
WILKIE, A. O. 2005. Bad bones, absent smell, selfish testes: the pleiotropic consequences of human FGF receptor mutations. *Cytokine Growth Factor Rev,* 16**,** 187-203.
WILLIAMS, S. V., HURST, C. D. & KNOWLES, M. A. 2013. Oncogenic FGFR3 gene fusions in bladder cancer. *Hum Mol Genet,* 22**,** 795-803.
WILSON, T. R., FRIDLYAND, J., YAN, Y., PENUEL, E., BURTON, L., CHAN, E., PENG, J., LIN, E., WANG, Y., SOSMAN, J., RIBAS, A., LI, J., MOFFAT, J., SUTHERLIN, D. P., KOEPPEN, H., MERCHANT, M., NEVE, R. & SETTLEMAN, J. 2012. Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature,* 487**,** 505-9.
WOHRLE, S., HENNINGER, C., BONNY, O., THUERY, A., BELUCH, N., HYNES, N. E., GUAGNANO, V., SELLERS, W. R., HOFMANN, F., KNEISSEL, M. & GRAUS PORTA, D. 2013. Pharmacological inhibition of fibroblast growth factor (FGF) receptor signaling ameliorates FGF23-mediated hypophosphatemic rickets. *J Bone Miner Res,* 28**,** 899-911.
WOLF, I., LEVANON-COHEN, S., BOSE, S., LIGUMSKY, H., SREDNI, B., KANETY, H., KURO-O, M., KARLAN, B., KAUFMAN, B., KOEFFLER, H. P. & RUBINEK, T. 2008. Klotho: a tumor suppressor and a modulator of the IGF-1 and FGF pathways in human breast cancer. *Oncogene,* 27**,** 7094-105.
WONG, A., LAMOTHE, B., LEE, A., SCHLESSINGER, J. & LAX, I. 2002. FRS2 alpha attenuates FGF receptor signaling by Grb2-mediated recruitment of the ubiquitin ligase Cbl. *Proc Natl Acad Sci U S A,* 99**,** 6684-9.
WONG, J. P., TODD, J. R., FINETTI, M. A., MCCARTHY, F., BRONCEL, M., VYSE, S., LUCZYNSKI, M. T., CROSIER, S., RYALL, K. A., HOLMES, K., PAYNE, L. S., DALEY, F., WAI, P., JENKS, A., TANOS, B., TAN, A. C., NATRAJAN, R. C., WILLIAMSON, D. & HUANG, P. H. 2016. Dual Targeting of PDGFRalpha and FGFR1 Displays Synergistic Efficacy in Malignant Rhabdoid Tumors. *Cell Rep,* 17**,** 1265-1275.
WRIGHT, J. D., AN, S. W., XIE, J., YOON, J., NISCHAN, N., KOHLER, J. J., OLIVER, N., LIM, C. & HUANG, C. L. 2017. Modeled structural basis for the recognition of alpha2-3-sialyllactose by soluble Klotho. *FASEB J*.
WU, A. L., COULTER, S., LIDDLE, C., WONG, A., EASTHAM-ANDERSON, J., FRENCH, D. M., PETERSON, A. S. & SONODA, J. 2011. FGF19 regulates cell proliferation, glucose and bile acid metabolism via FGFR4-dependent and independent pathways. *PLoS One,* 6**,** e17868.
WU, X., GE, H., GUPTE, J., WEISZMANN, J., SHIMAMOTO, G., STEVENS, J., HAWKINS, N., LEMON, B., SHEN, W., XU, J., VENIANT, M. M., LI, Y. S., LINDBERG, R., CHEN, J. L., TIAN, H. & LI, Y. 2007. Co-receptor requirements for fibroblast growth factor-19 signaling. *J Biol Chem,* 282**,** 29069-72.
WU, X., GE, H., LEMON, B., VONDERFECHT, S., WEISZMANN, J., HECHT, R., GUPTE, J., HAGER, T., WANG, Z., LINDBERG, R. & LI, Y. 2010. FGF19-induced hepatocyte proliferation is mediated through FGFR4 activation. *J Biol Chem,* 285**,** 5165-70.
WU, X., LEMON, B., LI, X., GUPTE, J., WEISZMANN, J., STEVENS, J., HAWKINS, N., SHEN, W., LINDBERG, R., CHEN, J. L., TIAN, H. & LI, Y. 2008. C-terminal tail of FGF19 determines its specificity toward Klotho co-receptors. *J Biol Chem,* 283**,** 33304-9.
WU, Y. M., SU, F., KALYANA-SUNDARAM, S., KHAZANOV, N., ATEEQ, B., CAO, X., LONIGRO, R. J., VATS, P., WANG, R., LIN, S. F., CHENG, A. J., KUNJU, L. P., SIDDIQUI, J., TOMLINS, S. A., WYNGAARD, P., SADIS, S., ROYCHOWDHURY, S., HUSSAIN, M. H., FENG, F. Y., ZALUPSKI, M. M., TALPAZ, M., PIENTA, K. J., RHODES, D. R., ROBINSON, D. R. & CHINNAIYAN, A. M. 2013. Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov,* 3**,** 636-47.
XIE, B., CHEN, J., LIU, B. & ZHAN, J. 2013. Klotho acts as a tumor suppressor in cancers. *Pathol Oncol Res,* 19**,** 611-7.

XIE, M. H., HOLCOMB, I., DEUEL, B., DOWD, P., HUANG, A., VAGTS, A., FOSTER, J., LIANG, J., BRUSH, J., GU, Q., HILLAN, K., GODDARD, A. & GURNEY, A. L. 1999. FGF-19, a novel fibroblast growth factor with unique specificity for FGFR4. *Cytokine,* 11**,** 729-35.

XIN, X., ABRAMS, T. J., HOLLENBACH, P. W., RENDAHL, K. G., TANG, Y., OEI, Y. A., EMBRY, M. G., SWINARSKI, D. E., GARRETT, E. N., PRYER, N. K., TRUDEL, S., JALLAL, B., MENDEL, D. B. & HEISE, C. C. 2006. CHIR-258 is efficacious in a newly developed fibroblast growth factor receptor 3-expressing orthotopic multiple myeloma model in mice. *Clin Cancer Res,* 12**,** 4908-15.

XU, A. M. & HUANG, P. H. 2010. Receptor tyrosine kinase coactivation networks in cancer. *Cancer Res,* 70**,** 3857-60.

XU, H., LEE, K. W. & GOLDFARB, M. 1998a. Novel recognition motif on fibroblast growth factor receptor mediates direct association and activation of SNT adapter proteins. *J Biol Chem,* 273**,** 17987-90.

XU, J., BUSSIERE, J., YIE, J., SICKMIER, A., AN, P., BELOUSKI, E., STANISLAUS, S. & WALKER, K. W. 2013. Polyethylene glycol modified FGF21 engineered to maximize potency and minimize vacuole formation. *Bioconjug Chem,* 24**,** 915-25.

XU, J., LLOYD, D. J., HALE, C., STANISLAUS, S., CHEN, M., SIVITS, G., VONDERFECHT, S., HECHT, R., LI, Y. S., LINDBERG, R. A., CHEN, J. L., JUNG, D. Y., ZHANG, Z., KO, H. J., KIM, J. K. & VENIANT, M. M. 2009. Fibroblast growth factor 21 reverses hepatic steatosis, increases energy expenditure, and improves insulin sensitivity in diet-induced obese mice. *Diabetes,* 58**,** 250-9.

XU, R., ORI, A., RUDD, T. R., UNIEWICZ, K. A., AHMED, Y. A., GUIMOND, S. E., SKIDMORE, M. A., SILIGARDI, G., YATES, E. A. & FERNIG, D. G. 2012. Diversification of the structural determinants of fibroblast growth factor-heparin interactions: implications for binding specificity. *J Biol Chem,* 287**,** 40061-73.

XU, X., WEINSTEIN, M., LI, C., NASKI, M., COHEN, R. I., ORNITZ, D. M., LEDER, P. & DENG, C. 1998b. Fibroblast growth factor receptor 2 (FGFR2)-mediated reciprocal regulation loop between FGF8 and FGF10 is essential for limb induction. *Development,* 125**,** 753-65.

YAMASHITA, T., YOSHIOKA, M. & ITOH, N. 2000. Identification of a novel fibroblast growth factor, FGF-23, preferentially expressed in the ventrolateral thalamic nucleus of the brain. *Biochem Biophys Res Commun,* 277**,** 494-8.

YAMAZAKI, Y., IMURA, A., URAKAWA, I., SHIMADA, T., MURAKAMI, J., AONO, Y., HASEGAWA, H., YAMASHITA, T., NAKATANI, K., SAITO, Y., OKAMOTO, N., KURUMATANI, N., NAMBA, N., KITAOKA, T., OZONO, K., SAKAI, T., HATAYA, H., ICHIKAWA, S., IMEL, E. A., ECONS, M. J. & NABESHIMA, Y. 2010. Establishment of sandwich ELISA for soluble alpha-Klotho measurement: Age-dependent change of soluble alpha-Klotho levels in healthy subjects. *Biochem Biophys Res Commun,* 398**,** 513-8.

YAMAZAKI, Y., TAMADA, T., KASAI, N., URAKAWA, I., AONO, Y., HASEGAWA, H., FUJITA, T., KUROKI, R., YAMASHITA, T., FUKUMOTO, S. & SHIMADA, T. 2008. Anti-FGF23 neutralizing antibodies show the physiological role and structural features of FGF23. *J Bone Miner Res,* 23**,** 1509-18.

YANG, C., JIN, C., LI, X., WANG, F., MCKEEHAN, W. L. & LUO, Y. 2012. Differential specificity of endocrine FGF19 and FGF21 to FGFR1 and FGFR4 in complex with KLB. *PLoS One,* 7**,** e33870.

YAYON, A., KLAGSBRUN, M., ESKO, J. D., LEDER, P. & ORNITZ, D. M. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell,* 64**,** 841-8.

YEH, B. K., IGARASHI, M., ELISEENKOVA, A. V., PLOTNIKOV, A. N., SHER, I., RON, D., AARONSON, S. A. & MOHAMMADI, M. 2003. Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. *Proc Natl Acad Sci U S A,* 100**,** 2266-71.

YIE, J., HECHT, R., PATEL, J., STEVENS, J., WANG, W., HAWKINS, N., STEAVENSON, S., SMITH, S., WINTERS, D., FISHER, S., CAI, L., BELOUSKI, E., CHEN, C., MICHAELS, M. L., LI, Y. S., LINDBERG, R., WANG, M., VENIANT, M. & XU, J. 2009. FGF21 N- and C-termini play different roles in receptor interaction and activation. *FEBS Lett,* 583**,** 19-24.

YIE, J., WANG, W., DENG, L., TAM, L. T., STEVENS, J., CHEN, M. M., LI, Y., XU, J., LINDBERG, R., HECHT, R., VENIANT, M., CHEN, C. & WANG, M. 2012. Understanding the physical interactions in the FGF21/FGFR/beta-Klotho complex:

structural requirements and implications in FGF21 signaling. *Chem Biol Drug Des,* 79**,** 398-410.
YLI-KAUHALUOMA, J. & TUOMINEN, R. K. 2011. Protein kinases as drug targets. *Curr Top Med Chem,* 11**,** 1304.
YU, X., IBRAHIMI, O. A., GOETZ, R., ZHANG, F., DAVIS, S. I., GARRINGER, H. J., LINHARDT, R. J., ORNITZ, D. M., MOHAMMADI, M. & WHITE, K. E. 2005. Analysis of the biochemical mechanisms for the endocrine actions of fibroblast growth factor-23. *Endocrinology,* 146**,** 4647-56.
ZEBISCH, K., VOIGT, V., WABITSCH, M. & BRANDSCH, M. 2012. Protocol for effective differentiation of 3T3-L1 cells to adipocytes. *Anal Biochem,* 425**,** 88-90.
ZHANG, J., GUPTE, J., GONG, Y., WEISZMANN, J., ZHANG, Y., LEE, K. J., RICHARDS, W. G. & LI, Y. 2017a. Chronic Over-expression of Fibroblast Growth Factor 21 Increases Bile Acid Biosynthesis by Opposing FGF15/19 Action. *EBioMedicine,* 15**,** 173-183.
ZHANG, X., GUREASKO, J., SHEN, K., COLE, P. A. & KURIYAN, J. 2006a. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell,* 125**,** 1137-49.
ZHANG, X., IBRAHIMI, O. A., OLSEN, S. K., UMEMORI, H., MOHAMMADI, M. & ORNITZ, D. M. 2006b. Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. *J Biol Chem,* 281**,** 15694-700.
ZHANG, X., STAPPENBECK, T. S., WHITE, A. C., LAVINE, K. J., GORDON, J. I. & ORNITZ, D. M. 2006c. Reciprocal epithelial-mesenchymal FGF signaling is required for cecal development. *Development,* 133**,** 173-80.
ZHANG, Y., HIRAISHI, Y., WANG, H., SUMI, K. S., HAYASHIDO, Y., TORATANI, S., KAN, M., SATO, J. D. & OKAMOTO, T. 2005. Constitutive activating mutation of the FGFR3b in oral squamous cell carcinomas. *Int J Cancer,* 117**,** 166-8.
ZHANG, Y., WANG, L., WU, Z., YU, X., DU, X. & LI, X. 2017b. The Expressions of Klotho Family Genes in Human Ocular Tissues and in Anterior Lens Capsules of Age-Related Cataract. *Curr Eye Res,* 42**,** 871-875.
ZHOU, H., HE, F., MENDELSOHN, C. L., TANG, M. S., HUANG, C. & WU, X. R. 2016. FGFR3b Extracellular Loop Mutation Lacks Tumorigenicity In Vivo but Collaborates with p53/pRB Deficiency to Induce High-grade Papillary Urothelial Carcinoma. *Sci Rep,* 6**,** 25596.
ZHOU, X. & WANG, X. 2015. Klotho: a novel biomarker for cancer. *J Cancer Res Clin Oncol,* 141**,** 961-9.
ZOU, X. & BLANK, M. 2017. Targeting p38 MAP kinase signaling in cancer through post-translational modifications. *Cancer Lett,* 384**,** 19-26.